

TOMATO SPOTTED WILT VIRUS

INS A1
Viruses in the *Tospovirus* genus infect a wide variety of plant species, particularly tobacco, peanut, vegetables and ornamental plants.

- 5 Two virus species, tomato spotted wilt virus (TSWV) and Impatiens necrotic spot virus (INSV) are recognized within the *Tospovirus* genus.

- Tomato Spotted Wilt Virus (TSWV) is unique among plant viruses in that the nucleic acid-protein complex is covered by a lipoprotein envelope and it is the only thrip transmitted virus. This virus has
- 10 recently been classified as the *Tospovirus* genus of the *Bunyaviridae* family. TSWV virions contain a 29K nucleocapsid protein ("NP" or "N"), two membrane-associated glycoproteins (58K and 78K) and a large 200K protein presumably for the viral transcriptase [see J. Gen. Virol. 71:2207 (1991); Virol. 56:12 (1973); and J. Gen. Virol. 36:267 (1977)].
- 15 The virus genome consists of three negative-strand (-) RNAs designated L RNA (8900 nucleotides), M RNA (5400 nucleotides) and S RNA (2900 nucleotides) [see J. Gen. Virol. 36:81 (1977); J. Gen. Virol. 53:12 (1981); and J. Gen. Virol. 70:3469 (1989)], each of which is encapsulated by the NP. The partial or full-length sequences of S RNAs from three TSWV
- 20 isolates reveals the presence of two open reading frames (ORF) with an ambisense gene arrangement [see J. Gen. Virol. 71:1 (1990) and J. Gen. Virol. 72:461 (1991)]. The larger open reading frame is located on the viral RNA strand and has the capacity to encode a 52K nonstructural protein. The smaller ORF is located on the viral complementary RNA
- 25 strand and is translated through a subgenomic RNA into the 29K NP.

- The ambisense coding strategy is also characteristic of the TSWV M RNA, with the open reading frames encoding the 58K and 78K membrane-associated glycoproteins. The TSWV L RNA has been sequenced to encode a large 200K protein presumably for the viral
- 30 transcriptase.

- Two TSWV serogroups, "L" and "I", have been identified and characterized based on serological analysis of the structural proteins and morphology of cytopathic structures [see J. Gen. Virol. 71:933 (1990) and Phytopathology 81:525 (1991)]. They have serologically
- 35 conserved G1 and G2 glycoproteins, but the NP of the "I" serogroup is

serologically distinct from that of the "L" serogroup. Comparison of the NP between the "L" and "I" serogroups has shown 62% and 67% identities at nucleotide and amino acid levels, respectively [see J. Gen. Virol. 72:2597 (1991)].

- 5 TSWV has a wide host range, infecting more than 360 plant species of 50 families and causes significant economic losses to vegetables and ornamental plants worldwide. The "L" serogroup has been found extensively in field crops such as vegetables and weeds, while the "I" serogroup has been largely confined to ornamental crops.
- 10 A cucurbit isolate has recently been identified [see Plant Disease 68:1006 (1984)] as a distinct isolate because it systemically infects watermelon and other cucurbits and its NP is serologically unrelated to that of either serogroup. Although the spread of the TSWV disease can sometimes be reduced by breeding resistant plants or using non-
- 15 genetic approaches, complete control of the disease by these conventional methods has generally proven to be difficult [see Plant Disease 73:375 (1989)].

- Since 1986, numerous reports have shown that transgenic plants with the coat protein (CP) gene of a virus are often resistant to
- 20 infection by that virus. This phenomenon is commonly referred to as coat protein-mediated protection (CPMP). The degree of protection ranges from delay in symptom expression to the absence of disease symptoms and virus accumulation. Two recent independent reports [see Biol. Technology 9:1363(1991) and Mol. Plant-Microbe Interact.
- 25 5:34 (1992)] showed that transgenic tobacco plants expressing the nucleocapsid protein (NP) gene of TSWV are resistant to infection by the homologous isolate. However, since TSWV is widespread with many biologically diverse isolates, it is very important to test the effectiveness of the transgenic plants to resist infections by different
- 30 TSWV isolates. The findings of the present invention expand on those of the previous reports by demonstrating that transgenic plants according to the present invention showed resistance to two heterologous isolates of the "L" serogroup and an isolate of the "I" serogroup. We also show that resistance to the two heterologous isolates of the "L"
- 35 serogroup was mainly found in plants accumulating very low, if any,

levels of NP, while transgenic plants that accumulated high levels of NP were resistant to the isolate of the "I" serogroup.

However, no resistance was observed to a Brazilian isolate, although the plants that accumulated high levels of the N protein did display a delay in symptom expression. This Brazilian isolate, designated TSWV-B has the N protein that was serologically distinct from the "L" and "I" serogroups and biologically differs from a curcubit isolate in that the TSWV-B does not systemically infect melons or squash. Therefore, one aspect of the present invention is to characterize the TSWV-B by cloning and sequencing of its S RNA and comparisons with the published sequences of other TSWV isolates.

Various aspects of the present invention will become readily apparent from the detailed description of the present invention including the following example, figures and data.

In the Figures;

Fig. 1 depicts the strategy for cloning the NP gene from viral RNA according to the present invention;

Fig. 2 depicts the ^{in-vivo} transient expression of the nucleocapsid protein (NP) gene of tomato spotted wilt virus according to the present invention in tobacco protoplasts;

Fig. 3 depicts the location of the sequenced cDNA clones in the TSWV-B S RNA according to the present invention;

Fig. 4 depicts a dendogram showing relationships among TSWV isolates according to the present invention;

Fig. 5 depicts the serological relationship of TSWV isolates described herein;

Fig. 6 depicts the correlation of the level of nucleocapsid protein (NP) accumulation in transgenic plants with the degree of resistance to TSWV isolates;

Fig. 7 depicts the TSWV-BL N coding sequences introduced into transgenic plants in accordance with one aspect of the present invention; and

Fig. 8 depicts the TSWV-BL half N gene fragments introduced into plants in accordance with one aspect of the present invention.

More specifically, figure 2 depicts transient expression of the NP gene in which the constructs were transferred into tobacco mesophyll protoplasts using polyethylene glycol (PEG). The transformed protoplasts were subsequently incubated for two days for the expression of the NP gene. Proteins were extracted from the protoplasts and tested for the NP by double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) using antibodies against the TSWV NP. NP⁻ and NP⁺ represent the protoplasts transformed with plasmids pBI525-NP⁻ and pBI525-NP⁺, respectively. Concentration of the antibodies for coating: 5 µg/ml; dilution of the enzyme conjugate: 1:250. Data were taken 30, 60 and 90 min. after addition of substrate.

In figure 3, the five overlapping cDNA clones are shown to scale below a S RNA map of TSWV-B. These clones were synthesized with random primers from double-stranded RNA isolated from *N. benthamiana* plants infected with TSWV-B.

In figure 4, the sequences were compared using the pileup program of the GCG Sequence analysis software package. Horizontal lines are proportional to the genetic distance while vertical lines are of arbitrary length and have no significance.

More specifically, in figure 5, *N. benthamiana* Domin. were infected with TSWV isolates [TSWV-BL (a lettuce isolate), Arkansas, 10W pakchoy (TSWV-10W), Begonia, and Brazil (TSWV-B)]. An infected leaf disc (0.05 gram) was ground in 12 ml of the enzyme conjugate buffer and analyzed by DAS-ELISA using antibodies raised against TSWV-BL ^{Virion} ~~viren~~ (BL ^{Virion} ~~viren~~), or the NP of TSWV-BL (BL-NP), or TSWV-I (I-NP). Concentration of antibodies for coating were 1 µg/ml; dilution of conjugates were 1:2000 for BL ^{Virion} ~~viren~~, 1:250 for BL-NP, and 1:1000 for I-NP. The results were taken after 10 minutes (BL), 50 minutes (BL-NP), or 30 minutes after adding substrate.

With regard to figure 6, transgenic plants were assayed in DAS-ELISA for NP accumulation with antibodies raised against the NP of TSWV-BL. Plants were read 150 min. after adding substrate and the transgenic plants were grouped into four categories: OD_{405nm} smaller than 0.050, OD_{405nm} between 0.050 to 0.200, OD_{405nm} between 0.200 to 0.400, and OD_{405nm} greater than 0.400. The OD_{405nm} readings of

- control NP (-) plants were from zero to 0.05. The same plants were challenged with either the Arkansas (Ark) and 10W pakchoy (10W) isolates or the Begonia isolate and the susceptibility of each plant was recorded about 12 days after inoculation. The results were pooled from
- 5 fifty-one R₁ NP (+) plants inoculated with the Arkansas and 10W pakchoy isolates and one hundred thirty-nine R₁ NP(+) plants inoculated with the Begonia isolate. Numbers above bars represent total numbers of R₁ NP(+) plants tested.

EXAMPLE I

10 Isolation of TSWV-BL RNAs:

- The TSWV-BL isolate was purified from *Datura stramonium* L. as follows: the infected tissues were ground in a Waring Blender for 45 sec with three volumes of a buffer (0.033 M KH₂PO₄, 0.067 M K₂HPO₄, 0.01 M Na₂SO₃). The homogenate was filtered through 4 layers of
- 15 cheesecloth moistened with the above buffer and centrifuged at 7,000 rpm for 15 min. The pellet was resuspended in an amount of 0.01 M Na₂SO₃ equal to the original weight of tissue and centrifuged again at 8,000 rpm for 15 min. After the supernatant was resuspended in an amount of 0.01 M Na₂SO₃ equal to 1/10 of the original tissue weight,
- 20 ~~the~~ The virus extract was centrifuged at 9,000 rpm for 15 min. and the supernatant was carefully loaded on a 10-40% sucrose step gradient made up in 0.01 M Na₂SO₃. After centrifugation at 23,000 rpm for 35 min., the virus zone (about 3 cm below meniscus) was collected and diluted with two volumes of 0.01 M Na₂SO₃. The semi-purified virus
- 25 was pelleted at 27,000 rpm for 55 min.

EXAMPLE II

Purification of TSWV and viral RNAs:

- The TSWV-BL isolate [see Plant Disease 74:154 (1990)] was purified from *Datura stramonium* L. as described in Example I. The
- 30 purified virus was resuspended in a solution of 0.04% of bentonite, 10 µg/ml of proteinase K, 0.1 M ammonium carbonate, 0.1% (w/v) of sodium diethyldithiocarbamate, 1 mM EDTA, and 1% (w/v) of sodium dodecyl sulfate (SDS), incubated at 65°C for 5 min., and immediately extracted from H₂O-saturated phenol, followed by another extraction

with chloroform/isoamyl alcohol (24:1). Viral RNAs were precipitated in 2.5 volumes of ethanol and dissolved in distilled H₂O.

EXAMPLE III

cDNA and PCR-based NP gene cloning:

5 The first strand cDNA was synthesized from purified TSWV-BL RNAs using random primers as described by Gubler and Hoffman [see Gene 25:263 (1983)]. The second strand was produced by treatment of the sample with RNase H/DNA polymerase. The resulting double-stranded cDNA sample was size-fractionated by sucrose gradient
10 centrifugation, methylated by EcoRI methylase, and EcoRI linkers were added. After digestion with EcoRI, the cDNA sample was ligated into the EcoRI site of pUC18, whose 5'-terminal phosphate groups were removed by treatment with calf intestinal alkaline phosphatase. *E. coli* DH5 α competent cells (Bethesda Research Laboratories) were
15 transformed and clones containing TSWV cDNA inserts were first selected by plating on agar plates containing 50 μ g/ml of ampicillin, IPTG, and X-gal. Plasmid DNAs from selected clones were isolated using an alkaline lysis procedure [see BRL Focus 11:7 (1989)], and the insert sizes were determined by EcoRI restriction enzyme digestion
20 followed by DNA transfer onto GeneScreen Plus nylon filters (DuPont). Plasmid clones that contained a TSWV-BL S RNA cDNA insert were identified as described below by hybridizing against a ³²P-labelled oligomer (AGCAGGCAAACTCGCAGAACTTGCY) ^(SEQ. ID. No. 1) complementary to the nucleotide sequence (GCAAGTTCTGCGAGTTTTGCCTGCT) ^(SEQ. ID. No. 2) of the TSWV-CPNH1 S RNA [see J. Gen. Virol. 71:001 (1990)]. Several clones were identified and analyzed on agarose gels to determine the insert sizes. The ~~clones~~ ^{clone} pTSWVS-23 was found to contain the largest cDNA insert, about 1.7 kb in length.

The full-length NP gene was obtained by the use of polymerase
30 chain reaction (PCR). First-strand cDNA synthesis was carried out at 37°C for 30 min. In a 20 μ l reaction mixture using oligomer primer JLS90-46 (5'→3') AGCTAACCATGGTTAAGCTCACTAAGGAAAGC ^(SEQ. ID. No. 3) (also used to synthesize the nucleocapsid gene of TSWV-10W) which is complementary to the S RNA in the 5' terminus of TSWV NP gene
35 (nucleotide positions 2751 to 2773 of the TSWV-CPNH1). The reaction

662207-ES252160

mixture contained 1.5 µg of viral RNAs, 1 µg of the oligomer primer, 0.2 mM of each dNTP, 1X PCR buffer (the GeneAmp kit, Perkin-Elmer-Cetus), 20U of RNAs In Ribonuclease Inhibitor (Promega), 2.5 mM of MgCl₂, and 25U of AMV reverse transcriptase (Promega Corporation). The reaction was terminated by heating at 95°C for 5 min. and cooled on ice. Then 10 µl of the cDNA/RNA hybrid was used to PCR-amplify the NP gene according to manufacturer's instructions (Perkin-Elmer-Cetus) using 1 µg each of oligomer primers JLS90-46 and JLS90-47 (5'→3'), AGCATTCCATGGTTAACAGACTAAGCAAGCAC (also used to synthesize the nucleotide gene of TSWV-10W), the latter oligomer being identical to the S RNA in the 3' noncoding region of the gene (nucleotide positions 1919 to 1938 of the TSWV-CPNH1). A typical PCR cycle was 1 min. at 92°C (denaturing), 1 min. at 50°C (annealing), and 2 min. at 72°C (polymerizing). The sample was directly loaded and separated on a 1.2% agarose gel. The separated NP gene fragment was extracted from the agarose gel, ethanol-precipitated and dissolved in 20 µl of distilled H₂O.

EXAMPLE IV

Construction of plant expression and transformation vectors.

The gel-isolated NP gene fragment from Example III was digested with the restriction enzyme *NcoI* in 50 µl of a reaction buffer [50 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, 0.1 M NaCl] at 37°C for 3 hours, and directly cloned into *NcoI*-digested plant expression vector pB1525. The resulting plasmids were identified and designated as pB1525-NP⁺ in the sense orientation relative to the cauliflower mosaic virus (CaMV) 35S promoter, and as pB1525-NP⁻ in the reverse orientation. The ability of this expression cassette to produce the NP was determined by transient expression of the NP gene in *Nicotiana tabacum* protoplasts, as described by Pang et al [see Gene 112:229 (1992)]. The expression cassette containing the NP gene was then excised from pB1525-NP⁺ by a partial digestion with *HindIII*/*EcoRI* (since the NP gene contains internal *HindIII* and *EcoRI* sites), and ligated into the plant transformation vector pBIN19 (Clontech Laboratories, Inc.) that had been cut with the same enzymes. The resulting vector, pBIN19-NP⁺ and the control plasmid pBIN19 were transferred to *A. tumefaciens* strain

LBA4404, using the procedure described by Holsters et al [see Mol. Gen. Genet. 163:181 (1978)].

5 Nucleotide sequence analyses of the inserts in clones pTSWV-23 and Pb1525-NP+ were determined using the dideoxynucleotide method, T7 polymerase (U.S. Biochemicals, SequenaseTM), and the double-stranded sequencing procedure described by Siemieniak et al [see Analyt. Biochem. 192:441 (1991)]. Nucleotide sequences were determined from both DNA strands and this information was compared with the published sequences of TSWV Isolates CPNH1 using computer
10 programs available from the Genetics Computer Group (GCG, Madison, WI).

Transient expression of the NP gene in tobacco protoplasts were also prepared. Plasmid DNAs for clones pTSWVS-23 and pUC18cpphas TSWV-NP (containing the PCR-engineered NP gene insert) were isolated
15 using the large scale alkaline method. The PCR-engineered NP gene insert was excised from clone pBIS25-NP+ by NcoI digestion to take advantage of the available flanking oligomer primers for sequencing. The expression cassette pUC18cpphas is similar to pUC18cpexp except that it utilizes the poly(A) addition signal derived from the *Phaseolus*
20 *vulgaris* seed storage gene phaseolin. These plasmid DNAs were subjected to two CsCl-ethidium bromide gradient bandings, using a Beckman Ti 70.1 fixed angle rotor. DNA sequences were obtained using dideoxynucleotides and the double-stranded plasmid DNA sequencing procedure described above. Nucleotide sequence reactions were
25 electrophoresed on one-meter long thermostated (55°C) sequencing gels and nucleotide sequence readings averaging about 750 bp were obtained. Nucleotide sequences were determined from both DNA strands of both cloned inserts to ensure accuracy. Nucleotide sequence information from the TSWV-BL S RNA isolate was compared as discussed below, with TSWV Isolates CPNH1 and L3 using computer programs (GCG,
30 Madison, WI).

The nucleotide and deduced amino acid sequences of cloned cDNA and PCR-engineered insert of TSWV-BL S RNA and their comparison with the nucleotide sequence of TSWV-CPNH1 S RNA are shown below.
35 The nucleotide sequence of the TSWV-BL S RNA clones pTSWVS-23

(TSWV-23) and pBI525-NP⁺ (TSWV-PCR) were obtained using the double-stranded dideoxynucleotide sequencing procedure of Slemieniak, and their sequences are compared with the relevant regions of the nucleotide sequence of the TSWV-CPNH1 S RNA reported in GeneBank Accession No. D00645. The nucleotide sequence of TSWV-CPNH1 S RNA has been reported by De Haan (1990) and is represented by the following sequence:

(SE01DN0.5)

	CAAGTTGAAA GCAACAACAG AACTGTAAAT TCCTCTGCAG TGAAATCTCT	50
	GCTCATGTCA GCAGAAAACA ACATCATGCC TAACTCTCAA GCTTCCACTG	100
10	ATTCTCATTT CAAGCTGAGC CTCTGGCTAA GGGTTCCAAA GGTTTTGAAG	150
	CAGGTTTCCA TTCAGAAATT GTTCAAGGTT GCAGGAGATG AAACAAACAA	200
	AACATTTTAT TTATCTATTG CCTGCATTCC AAACCATAAC AGTGTGTGAGA	250
	CAGCTTTAAA CATTACTGTT ATTTGCAAGC ATCAGCTCCC AATTGCGAAA	300
	TGCAAAGCTC CTTTGAATT ATCAATGATG TTTTCTGATT TAAAGGAGCC	350
15	TTACAACATT GTTCATGACC CTTTATACC CAAAGGATCG GTTCCAATGC	400
	TCTGGCTCGA AACTCACACA TCTTTGCACA AGTTCCTTGC AACTAACTTG	450
	CAAGAAGATG TAATCATCTA CACTTTGAAC AACCTTGAGC TAACTCCTGG	500
	AAAGTTAGAT TTAGGTGAAA GAACCTTGAA TTACAGTGAA GATGCGTACA	550
	AAAGGAAATA TTTCCTTTCA AAAACACTTG AATGTCTTCC ATCTAACACA	600
20	CAAACATATG CTTACTTAGA CAGCATCCAA ATCCCTTCAT GGAAGATAGA	650
	CTTTGCCAGA GGAGAAATTA AAATTTCTCC ACAATCTATT TCAGTTGCAA	700
	AATCTTTGTT AAAGCTTGAT TTAAGCGGGA TCAAAAAGAA AGAATCTAAG	750
	GTTAAGGAAG CGTATGCTTC AGGATCAAAA TAATCTTGCT TTGTCCAGCT	800
	TTTTCTAATT ATGTTATGTT TATTTTCTTT CTTTACTTAT AATTATTTCT	850
25	CIGTTTGTCA TCCTTTTCAA ATTCTCTCTG TCTAGTAGAA ACCATAAAAA	900
	CAAAAAATAA AAATGAAAAT AAAATTAAAA TAAAATAAAA TCAAAAAATG	1000
	AAATAAAAAC AACAAAAAAT TAAAAAACGA AAAACCAAAA AGACCCGAAA	1050
	GGGACCAATT TGGCCAAATT TGGGTTTTGT TTTTGTTTTT TGTTTTTTGT	1100
	TTTTTATTTT TTATTTTATT TTTATTTTAT TTTATTTTAA TTTTATTTTT	1150
30	ATTTTATTTA TTTTTTGT TTGTTGTTTT TGTATTTTAA TTATTTATTA	1200
	AGCACAACAC ACAGAAAGCA AACTTTAATT AACACACTT ATTAAAAATT	1250
	TAACACACTA AGCAAGCACA AGCAATAAAG ATAAAGAAAG CTTATATAT	1300
	TTATAGGCTT TTTTATAATT TAACTTACAG CTGCTTTCAA GCAAGTTCTG	1350
	CGAGTTTTCG CTGCTTTTAA ACCCOGAACA TTTCATAGAA CTTGTTAAGA	1400
35	GTTCACATGT AATGTTCCAT AGCAACACTC CCTTAGCAT TAGGATTGCT	1450

66207-02-21-60

SUB #3
CONT

5 GGAGCTAAGT ATAGCAGCAT ACTCTTTOCC CTCTTCACC TGATCTTCAT 1500
 TCATTTCAAA TGCTTTGCTT TTCAGCACAG TGCAAACITT TCCTAAGGCT 1550
 TCCTTGGTGT CATACTTCTT TGGGTCGATC CCGAGGTCTT TGTATTTTGC 1600
 ATCCTGATAT ATAGCCAAGA CAACACTGAT CATCTCAAAG CTATCAACTG 1650
 AAGCAATAAG AGGTAAGCTA CCTOCCAGCA TTATGGCAAG TCTCACAGAC 1700
 TTTGCATCAT OGAGAGGTAA TCCATAGGCT TGAATCAAAG GATGGGAAGC 1750
 AATCTTAGAT TTGATAGTAT TGAGATTCTC AGAATTOCCA GTTCTTTCAA 1800
 CAAGCCTGAC OCTGATCAAG CTATCAAGCC TTCTGAAGGT CATGTCAGTG 1850
 CCTCCAATCC TGTCGAAGT TTCTTTTATG GTAATTTTAC CAAAAGTAAA 1900
 10 ATCGCTTTGC TTAATAACCT TCATTATGCT CTGACGATTG TTTAGGAATG 1950
 TCAGACATGA AATAAGCTC ATCTTCTTGA TCTGGTGGAT GTTTTCCAGA 2000
 CAAAAGTCT TGAAGTTGAA TGCTACCAGA TTCTGATCTT OCTCAAATC 2050
 AAGGCTTTTG CCTTGIGTCA ACAAGCAAC AATGCTTTCC TTAGTGAGCT 2100
 TAACTTAGA CATGATGATC GTAAAAGTTG TTATAGCTTT GACOGTATGT 2150
 15 AACTCAAGGT GCGAAAGTGC AACTCTGTAT CCGCAGTCG TTTCTTAGGT 2200
 TCTTAATGIG ATGATTGTGA AGACTGAGTG TTAACGTATG AACACAAAAT 2250
 TGACAAGATT GCTCT 2265

20 The incomplete deduced amino acid sequence of the nonstructural protein gene on TSWV-CPNH1 S RNA is provided below beginning with nucleic acid at position 1 and ending with the nucleic acid codon ending at position 783.
 (SEQ. ID No. 7)

SUB #4
25

Gln Val Glu Ser Asn Asn Arg Thr Val Asn Ser Leu Ala Val Lys
 5 10 15
 Ser Leu Leu Met Ser Ala Glu Asn Asn Ile Met Pro Asn Ser Gln
 20 25 30
 Ala Ser Thr Asp Ser His Phe Lys Leu Ser Leu Trp Leu Arg Val
 35 40 45
 Pro Lys Val Leu Lys Gln Val Ser Ile Gln Lys Leu Phe Lys Val
 50 55 60
 30 Ala Gly Asp Glu Thr Asn Lys Thr Phe Tyr Leu Ser Ile Ala Cys
 65 70 75
 Ile Pro Asn His Asn Ser Val Glu Thr Ala Leu Asn Ile Thr Val
 80 85 90
 Ile Cys Lys His Gln Leu Pro Ile Arg Lys Cys Lys Ala Pro Phe
 95 100 105
 35 Glu Leu Ser Met Met Phe Ser Asp Leu Lys Glu Pro Tyr Asn Ile
 110 115 120
 Val His Asp Pro Ser Tyr Pro Lys Gly Ser Val Pro Met Leu Trp
 125 130 135

Leu Glu Thr His Thr Ser Leu His Lys Phe Phe Ala Thr Asn Leu
 140 145 150
 Gln Glu Asp Val Ile Ile Tyr Thr Leu Asn Asn Leu Glu Leu Thr
 155 160 165
 5 Pro Gly Lys Leu Asp Leu Gly Glu Arg Thr Leu ~~Asn Tyr Ser~~ Glu
 170 175 180
 Asp Ala Tyr Lys Arg Asp Tyr Phe Leu Ser Lys Thr Leu Glu Cys
 185 190 195
 Leu Pro Ser Asn Thr Gln Thr Met Ser Tyr Leu Asp Ser Ile Gln
 200 205 210
 Ile Pro Ser Trp Lys Ile Asp Phe Ala Arg Gly Glu Ile Lys Ile
 215 220 225
 Ser Pro Gln Ser Ile Ser Val Ala Lys Ser Leu Leu Lys Leu Asp
 230 235 240
 15 Leu Ser Gly Ile Lys Lys Lys Glu Ser Lys Val Lys Glu Ala Tyr
 245 250 255
 Ala Ser Gly Ser Lys
 260

20 The nucleotide sequence for TSWV-23 depicted below compares closely with the TWSV sequence given above, and contains one-half of the nonstructural gene and one half of the nucleocapsid protein gene. (SEQ ID NO. 6)

AAATTCTCTT GCAGTGAAAT CTCTGCTCAT GTTAGCAGAA AACAAATCA 50
 TGCCTAACTC TCAAGCTTTT GTCAAAGCTT CTACTGATTG TAATTTCAAG 100
 CTGAGCCTCT GGCTAAGGGT TCCAAAGGTT TTGAAGCAGA TTTCCATTCA 150
 25 GAAATTGTTT AAGGTTGCAG GAGATGAAAC AAATAAAACA TTTTATTAT 200
 CTATTGCCTG CATTCCAAAC CATAACAGTG TTGAGACAGC TTTAAACATT 250
 ACTGTTATTT GCAAGCATCA GCTCCCAATT CGTAAATGTA AAACCTCTTT 300
 TGAATTATCA ATGATGTTTT CTGATTTAAA GGAGCCTTAC AACATTATTC 350
 ATGATCCTTC ATATCCCCAA AGGATTGTTT ATGCTCTGCT TGAAACTCAC 400
 30 ACATCTTTTG CACAAGTTCT TTGCAACAAC TTGCAAGAAG ATGIGATCAT 450
 CTACACCTTG AACAAACCATG AGCTAACTCC TGGAAAGTGA GATTTAGGTG 500
 AAATAACITT GAATTACAAT GAAGACGCCT ACAAAGGAA ATATTTCTTT 550
 TCAAAAACAC TTGAATGTCT TCCATCTAAC ATACAACTA TGCTTTATTT 600
 AGACAGCATC CAAATCCCTT CCTGGAAGAT AGACTTTGCC AGGGGAGAAA 650
 35 TTAATAATTC TCCACAATCT ATTTTCAGTTG CAAATCTTT GTTAAATCTT 700
 GATTTAAGCG GGATTAAAAA GAAAGAATCT AAGATTAAAG AAGCATATGC 750
 TTCAGGATCA AAATGATCTT GCTGTGTCCA GCTTTTTCTA ATTATGTTAT 800
 GTTATTTTTC TTCTTTTACT TATAATTATT TTTCTGTTTG TCATTTCTTT 850
 CAAATTCCTC CTGCTAGTGA GAAACCATAA AAACAAAAAT AAAAATAAAA 900

TAAAATCAAA ATAAAATAAA AATCAAAAAA TGAAATAAAA GCAACAAAAA 950
 AATTAAAAAA CAAAAAACCA AAAAAGATCC CGAAAGGACA ATTTTGGCCA 1000
 AATTGGGGT TTGTTTTGT TTTTGITTT TTTGTTTTT GTTTTATTT 1050
 TTATTTTAT TTTTATTTT ATTTTATTT ATTTATGTT TTTGTTGTT 1100
 5 TIGTTATTT GTTATTTATT AAGCACAACA CACAGAAAGCA AACTTTAAT 1150
 TAAACACACT TATTTAAAAT TTAACACACT AAGCAAGCACA AACAATAAA 1200
 GATAAAGAAA GCITTATATA TTTATAGGCT TTTTATAAT TTAACCTACA 1250
 GCTGCTTTTA AGCAAGTTCT GIGAGTTTTG CCGTTTTTTT AACCCCAAAC 1300
 ATTCATAGA ACTTGTTAAG GGTTTCACTG TAATGTTCCA TAGCAATACT 1350
 10 TCCTTTAGCA TTAGGATTGC TGGAGCTAAG TATAGCAGCA TACTCTTTCC 1400
 CCTTCTTAC CTGATCTTCA TTCAATTCAA ATGCTTTTCT TTTCAGCACA 1450
 GIGCAAACCT TTOCTAAGGC TTCCCTGGTG TCATACTTCT TTGGGTGAT 1500
 CCGAGATCC TTGTATTTTG CATCCTGATA TATAGCCAAG ACAACACTGA 1550
 TCATCTCAA GCTATCAACT GAAGCAATAA GAGGTAAGCT ACCTCCAGC 1600
 15 ATTATGGCAA GCCTCACAGA CTTTGCATCA TCAAGAGGTA ATCCATAGGC 1650
 TTGAATCAAA GGGTGGGAAG CAATCTTAGA TTTGATAGTA TTGAGATTCT 1700
 CAGAATCC 1709

The nucleic acid sequence for TSWV-PCR according to the present
 invention as depicted below also compares closely with the TSWV^(SEQ. ID NO. 8)
 20 sequence given above and covers the whole nucleocapsid protein gene.
 TTAACACACT AAGCAAGCAC AAACAATAAA GATAAAGAAA GCITTATATA 50
 TTTATAGGCT TTTTATAAT TTAACCTACA GCTGCTTTTA AGCAAGTTCT 100
 GIGAGTTTTG CCGTTTTTTT AACCCCAAAC ATTCATAGA ACTTGTTAAG 150
 GGTTTCACTG TAATGTTCCA TAGCAATACT TCCTTTAGCA TTAGGATTGC 200
 25 TGGAGCTAAG TATAGCAGCA TACTCTTTCC CCTTCTTAC CTGATCTTCA 250
 TTCAATTCAA ATGCTTTTCT TTTCAGCACA GIGCAAACCT TTOCTAAGGC 300
 TTCCCTGGTG TCATACTTCT TTGGGTGAT CCGAGATCC TTGTATTTTG 350
 CATCCTGATA TATAGCCAAG ACAACACTGA TCATCTCAA GCTATCAACT 400
 GAAGCAATAA GAGGTAAGCT ACCTCCAGC ATTATGGCAA GCCTCACAGA 450
 30 CTTTGCATCA TCAAGAGGTA ATCCATAGGC TTGACTCAA GGGTGGGAAG 500
 CAATCTTAGA TTTGATAGTA TTGAGATTCT CAGAATCCC AGTTTCTCA 550
 ACAAGCCIGA CCTGATCAA GCTATCAAGC CTTCTGAAG TCATGTCAGT 600
 GGCTCAATC CTGCTGAAG TTTCTTTAT GTTAATTTTA CAAAAGTAA 650
 AATCGCTTG CTAAATAAOC TTCAATTATG TCTGACGATT CTTCAGGAAT 700
 35 GTCAGACATG AAATAATGCT CATCTTTTG ATCTGGTCAA GGTTTCCAG 750

ACAAAAAGTC TTGAAGTTGA ATGCTACCAG ATTCTGATCT TCCTCAAAC 800
 CAAGGTCITTT GCCTTGIGIC AACAAAGCAA CAATGCTTTC CTTAGTGAGC 850
 TTAACCAT 858

Together the cloned TSWV-23 insert overlaps the TSWV-PCR
 5 insert, and together they represent the 2028 nucleotides of the TSWV-
 BL S RNA according to the present invention. This 2028 nucleotide
 sequence according to the present invention contains a part of the
 nonstructural gene and whole nucleocapsid protein gene. The combined
 sequence is: (SEQ ID NO. 9)

10 AAATTCICTT GCAGTGAAAT CTCIGCTCAT GTTAGCAGAA AACACATCA 50
 TGCCIAACTC TCAAGCTTTT GTCAAAGCTT CTA CTGATTTC TAATTTCAAG 100
 CTGAGCCTCT GGCTAAGGGT TOCAAAGGTT TTGAAGCAGA TTTCATTCA 150
 GAAATTGTTT AAGGTTGCAG GAGATGAAAC AAATAAAACA TTTTATTTAT 200
 CTATTGCTTG CATTCCAAAC CATAACAGTG TTGAGACAGC TTAAACATT 250
 15 ACTGTTATTT GCAAGCATCA GCTCCCAATT CGTAAATGTA AAATTCCTTT 300
 TGAATTATCA ATGATGTTTT CTGATTTAAA GGAGCCTTAC AACATTATTC 350
 ATGATCCTTC ATATCCCOCA AGGATTGTTT ATGCTCTGCT TGAAACTCAC 400
 ACATCTTTTG CACAAGTTCT TTGCAACAAC TTGCAAGAAG ATGTGATCAT 450
 CTACACCTTG AACCAACCATG AGCTAACTCC TGGAAAGTTA GATTTAGGTG 500
 20 AAATAACTTT GAATTACAAT GAAGACGCTT ACAAAGGAA ATATTTCCCTT 550
 TCAAAACAC TTGAATGTCT TOCATCTAAC ATACAAACTA TGTCTTATTT 600
 AGACAGCATC CAAATCCCTT CCTGGAAGAT AGACTTTGCC AGGGGAGAAA 650
 TAAAATTTT TOCACAATCT ATTTGAGTTG CAAAATCTTT GTTAAATCTT 700
 GATTTAAGCG GGATTAAAAA GAAAGAATCT AAGATTAAAG AAGCATATGC 750
 25 TTCAGGATCA AAATGATCTT GCTGIGTCCA GCTTTTCTA ATTATGTTAT 800
 GTTATTTTTC TTCTTTTACT TATAATTATT TTCTGTTTG TCATTTCTTT 850
 CAAATTCCTC CTGCTTAGTA GAAACCATAA AAACAAAAT AAAAATAAAA 900
 TAAAATCAAA ATAAAATAAA AATCAAAAAA TGAAATAAAA GCAACAAAAA 950
 AATTAAAAAA CAAAAACCA AAAAGATCC CGAAAGGACA ATTTTGGOCA 1000
 30 AATTGCGGGT TTGTTTTTGT TTTTGTFTT TTGTTTTTTT GTTTTATTT 1050
 TTATTTTAT TTTTATTTT ATTTATTTT ATTTATGTT TTGTTGTTT 1100
 TTGTTATTT GTTATTATT AAGCACAACA CACAGAAAGC AAATTTAAT 1150
 TAAACACACT TATTTAAAAT TTAACACACT AAGCAAGCAC AAACAATAAA 1200
 GATAAAGAAA GCTTTATATA TTTATAGGCT TTTTATAAT TTAATTTACA 1250
 35 GCTGCTTTTA AGCAAGTTCT GTGAGTTTTG CCTGTTTTTT AACCCCAAAC 1300

5 ATTTCATAGA ACTTGTTAAG GGTTTCACTG TAATGTTTCA TAGCAATACT 1350
 TCCTTTAGCA TTAGGATTGC TGGAGCTAAG TATAGCAGCA TACTCTTTCC 1400
 CCTTCTTCAC CTGATCTTCA TTCATTTCAA ATGCTTTTCT TTTCAGCACA 1450
 GTGCAAACTT TTCTAAGGC TTCCCTGGIG TCATACTTCT TTGGGTGCA 1500
 CCGGAGATCC TTGTATTTTG CATCTGATA TATAGCCAAG ACAACACTGA 1550
 TCATCTCAA GCTATCAACT GAAGCAATAA GAGGTAAGCT AACTCCAGC 1600
 ATTATGGCAA GCTCACAGA CTTTGCATCA TCAAGAGGTA ATCCATAGGC 1650
 TTGACTCAA GGGTGGGAAG CAATCTTAGA TTTGATAGTA TTGAGATTCT 1700
 CAGAATTCCC AGTTTCTCA ACAAGCCTGA CCTGATCAA GCTATCAAGC 1750
 10 CTCTGAAGG TCATGTCAGT GGCTCAATC CTGTCTGAAG TTTTCTTTAT 1800
 GGTAATTTTA CAAAAGTAA AATCGCTTTG CTTAATAACC TTCAATTATGC 1850
 TCIGACGATT CTTCAGGAAT GTCAGACATG AAATAATGCT CATCTTTTGT 1900
 ATCTGGTCAA GGTTTTCAG AAAAAAGTC TTGAAGTTGA ATGCTACCAG 1950
 ATTCTGATCT TCTCAAACCT CAAGGTCCTT GCCTTGTGTC AACAAAGCAA 2000
 15 CAATGCTTTC CTTAGTGAGC TTAACCAT 2028

This comparison showed that cDNA insert of clone pTSWVS-23 included about 760 bp of the 52 K protein viral component gene, the complete intergenic region (492 bp), and 450 bp of the NP gene (about half of the NP gene). This cloned insert had its 3'-end located exactly at an EcoRI recognition site, which suggested incomplete EcoRI methylation during the cDNA cloning procedure. Although this clone did not contain the complete TSWV-BL NP gene, its sequence was of considerable importance since it had a 450 bp overlap with the sequence of the PCR-engineered NP gene (a total of 2028 bp of the TSWV-BL S RNA is presented in the nucleotide sequence for TSWV). The sequence comparison between this TSWV-BL PCR-engineered and TSWV-CPNH1 NP genes revealed a total of 21 nucleotide differences (2.7%), eight of which encode amino acid replacements (3.1%). Since this PCR engineered NP gene was obtained using Taq polymerase, which is known to incorporate mutations, it is possible that some of these differences were introduced during PCR amplification. However, 15 of these nucleotide differences were located within the overlapping region between the TSWV-BL cDNA and PCR clones, and all but one of these nucleotide differences (position 1702 of TSWV; position 485 of TSWV-PCR)) are shared by both TSWV-BL S RNA derived clones. This

comparison clearly showed that the PCR amplification did not contribute greatly, if at all, to the difference between the nucleotide sequences of these two cloned NP gene regions. The nucleotide difference at position 1702 resulted in the amino acid replacement of Ile with Ser, and even this difference could be due to the lack of homogeneity within the TSWV-BL isolate.

EXAMPLE V

***Agrobacterium*-mediated transformation:**

Leaf discs of *Nicotiana tabacum* var *Havana* cv 423 were inoculated with the *Agrobacterium* strain LBA4404 (ClonTech) containing the vector pBIN19-NP⁺ or the control plasmid pBIN19, by soaking overnight in a liquid culture of the *Agrobacterium*, and the inoculated leaf discs were incubated on non-selective MS medium for 3 days. [see Science 227:1229 (1985)]. Transformed cells were selected and regenerated in MS medium containing 300 µg/ml kanamycin and 500 µg/ml carbenicillin for shoot regeneration. Roots were induced after transfer of plantlets to hormone-free medium. Rooted transformants were transferred to soil and grown under greenhouse conditions. The MS medium contains full strength MS salt (Sigma), 30 g/l sucrose, 1 mg/l BA and 1 ml of B5 vitamins [1 mg/ml Nicotinic acid, 10 mg/ml Thiamine (HCl), 1 mg/ml Pyridoxine (HCl), 100 mg/ml Myo-Inositol]. Transgenic plants were self-pollinated and seeds were selectively germinated on kanamycin medium.

EXAMPLE VI

Serological detection of proteins:

Double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA) was used to detect the expression of NP gene in transgenic plants with polyclonal antibodies against the TSWV-BL NP. Each sample was prepared by grinding a leaf disc (about 0.05 g) from the top second leaf of the plant in 3 ml of an enzyme conjugate buffer [phosphate-buffered saline, 0.05% Tween 20, 2% polyvinylpyrrolidone 40, and 0.2% ovalbumin]. For tobacco protoplasts, the cell extracts after centrifugation were directly used for the assay. A ten- and three-fold dilutions of the samples from both transgenic plants and tobacco protoplasts were made just before DAS-ELISA.

For Western blots, a leaf disc (about 0.05 g) was ground in 0.25 ml of 2X SDS/sample buffer (0.126 M Tris buffer, 20% glycerol, 2% SDS, 2% 2-mercaptoethanol, and 0.01 mg/ml bromphenol blue). The homogenates were centrifuged and the supernatants were boiled before loading. Proteins (10-20 µl sample/lane) were separated and blotted onto a membrane. The membrane was then processed following the manufacturer's Immunoselct kit instruction manual (Gibco BRL Life Technologies Inc.). Antibodies to the whole virion were preabsorbed with cell extracts from health tobacco plants [See Plant Disease 70:501 (1986)], and were used in Western blot at a concentration of 2 µg/ml.

Serological reactions of TSWV isolates (TSWV-BL, Arkansas, 10W pakchoy, Begonia or Brazil) were assayed in DAS-ELISA using antibodies raised against TSWV-BL virion, or the NP of TSWV-BL or TSWV-I.

EXAMPLE VII

Inoculation of transgenic plants with TSWV Isolates.

Inocula were prepared by infecting *Nicotiana benthamiana* Domin. with different TSWV Isolates and grinding infected leaves (0.5 g) of *N. benthamiana* plants (1 to 2 weeks after inoculation) in 15 ml. of a buffer (0.033 M KH₂PO₄, 0.067 M K₂HPO₄ and 0.01 M Na₂SO₃). The inoculum extracts were immediately rubbed on corundum-dusted leaves of transgenic plants and the inoculated leaves were subsequently rinsed with H₂O. Because TSWV is highly unstable ~~in vitro~~ after grinding, each batch of inoculum was used to first inoculate NP(+) plants containing the NP gene; the last inoculated plants of each inoculum were always control NP(-) plants containing the vector sequence alone to assure that a particular virus inoculum was still infective at the end of inoculation.

Data on local lesions and systemic infections were taken 7-15 days after inoculation and expressed in the following table as the number of plants systemically infected over the number of plants inoculated, except where indicated. In this table, the data collected under "ELISA" is the data of R₀ lines from which the R₁ plants were ^{and} derived; the Begonia isolate induced local lesions on the R₁ plants, ~~and~~ the resistance was expressed as the number of plants producing local

Insert
AS

lesions over the number of plants inoculated; and NT indicates that there was no test.

Reactions of R₁ plants expressing the nucleocapsid protein (NP) gene of tomato spotted wilt virus (TSWV) to inoculation with TSWV isolates.

5

Reactions to TSWV Isolates

	ELISA: (R0 pl.)	BL	Arkansas	10W Pakchoy	Begonia	Brazil
<u>R₀ line</u>						
10 NP(+)	2 0.015	0/20	4/25	3/24	29/40	36/36
NP(+)	4 0.386	6/30	21/23	18/21	9/48	42/42
NP(+)	9 0.327	0/20	NT	20/20	—	—
NP(+)	14 0.040	0/20	—	9/20	8/18	18/18
NP(+)	21 0.042	0/15	5/15	3/15	2/4	6/6
15 NP(+)	22 0.142	0/20	—	15/20	31/36	36/36
NP(+)	23 0.317	0/20	—	16/20	—	—
NP(-)	-	42/42	24/24	62/62	66/66	54/54

As described above, the isolation of the TSWV-BL NP gene, which resides in the S RNA component of TSWV, was approached using two strategies. The cDNA cloning strategy yielded several clones containing cDNA inserts derived from TSWV-BL S RNA, as identified by hybridization against an oligomer probe complementary to the TSWV-CPNH1 S RNA. Clone pTSWVS-23 contained the longest insert, which mapped at about 1.7 kb in length. The second strategy utilized the published sequence of TSWV-CPNH1 S RNA and PCR to amplify and engineer the NP gene for expression directly from total TSWV-BL RNA. Oligomer primers JLS90-46 and -47 were synthesized, with JLS90-46 being complementary to the S RNA in the 5'-coding region of the NP gene (positions 2051-2073 of the TSWV-CPNH1) while JLS90-47 being of the 3'-noncoding region of the NP gene (positions 1218 to 1237 of the TSWV-CPNH1). Both of the primers contain the recognition site for the restriction enzyme *Nco*I for subsequent cloning, and the primer JLS90-46 has a plant consensus translation initiation codon sequence (AAXXATGG), which upon amplification was expected to fuse the translation initiation codon to the third codon (GTT) of the NP gene. Fusion of the translation initiation codon to the third codon of the

TSWV-BL NP gene was done to preserve the *NcoI* recognition site while not incorporating any amino acid codons. Thus, expression of the PCR-engineered TSWV NP gene would yield a TSWV-BL NP that was two amino acids (Ser-Lys) shorter at the N-terminus than the native NP.

5 This specifically-amplified DNA fragment, of about 850 bp, was digested with *NcoI* and cloned into the plant expression vector pB1525. The orientation of the TSWV-BL NP gene with respect to the CaMV 35S promoter was determined by restriction enzyme site mapping (*EcoRI*, *HindIII*, *AvaI* and *AluNI*). Several clones were isolated that contain the
10 insert in the proper orientation (pB1525-NP⁺) and others that contain the insert in the opposite orientation (pB1525-NP⁻). This restriction enzyme site mapping data also showed that the inserts of clones pB1525-NP⁺ contained restriction enzyme sites that were identical to those found in the TSWV-CPNH1 NP gene. The expression of TSWV-BL
15 NP gene was thus controlled by a double CaMV 35S promoter fused to the 5'-untranslated leader sequence of alfalfa mosaic virus (ALMV) of the expression vector pB1525. Expression vectors that utilize the stacked double CaMV 35S promoter elements yield higher levels of mRNA transcription than similar vectors that utilize a single 35S
20 promoter element.

Three pB1525-NP⁺ clones were transiently expressed in tobacco protoplasts to confirm that the amplified DNA fragment encoded the NP. To achieve this, the clones were transferred into tobacco protoplasts by the PEG method, and after two days of incubation the expressed NP
25 was detected by DAS-ELISA using antibodies against the whole TSWV-BL virion. High levels of NP were produced in tobacco protoplasts harboring the NP gene in plasmid pB1525-NP⁺; while no NP was detected in tobacco protoplasts transformed with the antisense NP sequence (pB1525-NP⁻).

30 As described previously, the PCR-engineered insert of clone pB1525-NP⁺ and ^{the} cDNA insert of the clone pTSWV-23 were subjected to double stranded sequencing. The sequence analysis of the cDNA and the PCR clones revealed inserts of 1.71 kb and 865 bp, respectively which, when compared with the sequence TSWV-CPNH1 S RNA, shows
35 that cDNA insert of clone pTSWV-23 includes about 760 bp of the 52 K

protein viral component gene, the complete intergenic region (492 bp), and 450 bp of the NP gene (about one-half of the gene). This cloned insert has its 3'-end located exactly at an *EcoRI* recognition site suggesting incomplete *EcoRI* methylation during the cDNA cloning procedure. Although this clone does not contain the complete TSWV-BL NP gene, its sequence is of considerable importance since it has a 450 bp overlap with the sequence of the PCR-engineered NP gene. The sequence comparison between this TSWV-BL PCR-engineered and TSWV-CPNH1 NP genes reveals a total of 21 nucleotide differences (2.7%), eight of which encode amino acid replacements (3.1%). Since this PCR-engineered NP gene was obtained using *Taq* polymerase, which is known to incorporate mutations, it is possible that some of these differences were introduced during PCR amplification. However, 15 of these nucleotide differences are located within the overlapping region between the TSWV-BL cDNA and PCR clones, and all but one of these differences (position 1702) are present in both TSWV-BL S RNA derived clones. This comparison clearly shows that the PCR amplification did not contribute greatly, if at all, to the difference between the nucleotide sequences of these two NP genes. The nucleotide difference at position 1702 results in the amino acid replacement of Ile with Ser, and even this difference could be due to the lack of homogeneity within the TSWV-BL isolate.

The possibility that the nucleotide differences can be attributed to divergence among the TSWV isolates is also supported by comparisons with other sequenced regions among TSWV-CPNH1, TSWV-~~TSWV-BL~~ L3, and ~~TSWV-BL~~ S RNAs. These comparisons are tabulated below:

Percent nucleotide and amino acid sequence differences for the comparison of TSWV S RNA component from isolates CPNH1, L3 and BL^a

Comparison	<u>52 K Protein Gene</u>		<u>Intergenic</u>	<u>NP Gene</u>	
	Nucleotide	Amino Acid	Nucleotide	Nucleotide	Amino Acid
CPNH1/L3	68/1396 ^b (4.9) ^c	49/464(10.6)	46/511(9.0)	24/777(3.1)	4/258(1.6)
CPNH1/BL	21/758(4.1)	23/251(9.2)	26/496(5.2)	19/765(2.5)	8/255(3.1)
L3/BL	38/765(5.0)	20/254(7.9)	38/498(7.6)	19/767(2.5)	4/255(1.6)

a Comparisons are made using the sequence information available from the particular component region of TSWV-BL. The comparison for the TSWV-BL NP gene includes the combined sequence information from the cDNA clone, pTSWVS-23 and PCR-engineered insert.

5 b Comparison numbers are total differences (nucleotides or amino acids) divided by total number of positions (nucleotides or amino acids) compared. For both nucleotide and amino acid calculation gaps, regardless of length, were counted as one mismatch.

c Numbers in parentheses are percentages.

10 The nucleotide sequence of the NP genes from the CPNH1 and L3 isolates differ from each other by 3.1% and from the BL isolate by nearly a similar degree (2.5%). However, the NP amino acid sequences between CPNH1 and BL isolates differ by a considerably larger amount than they differ between the L3 and BL or CPNH1 and L3 isolates. The results tabulated above also reveal that the NP gene region of these
15 TSWV isolates is subject to a higher degree of selective pressure than the 52 K protein as the differences among the amino acid sequences of the 52 K protein range between 7.9 to 10.6%, more than twice that found for the amino acid sequence of the NPs. Nucleotide sequence
20 divergence is highest among the intergenic regions, indicating that this region is subject to less selective pressure than either genetic region.

The presence of NP gene sequences in transgenic plants was first confirmed by PCR analysis. A NP DNA fragment of about 800 bp was specifically amplified from the total DNAs of transgenic NP(+) plants
25 using the primers homologous to sequences flanking the NP gene, whereas no corresponding fragment was detected in control NP(-) plants. Expression of the NP gene was assayed in each R₀ transgenic plant by DAS-ELISA, and the results are presented in the following table:

Reactions of R0 transgenic plants expressing the nucleocapsid protein (NP) gene of tomato spotted wilt virus (TSWV) to inoculation with TSWV-BL isolate

plant age	R0 clone	ELISA ^a	Lesions/leaf ^b	NP(+):NP(-) ^c
7-8 leaves:				
5	NP(+) ¹	0.374	7 (199)	1:28
	NP(+) ²	0.015	0 (199)	0:199
	NP(+) ³	0.407	23 (102)	1:4
	NP(+) ⁴	0.386	2 (102)	1:51
	NP(+) ⁵	0.023	0 (124)	0:124
10	NP(+) ⁶	0.197	35 (325)	1:9
	NP(+) ⁷	0.124	1 (325)	1:325
9-10 leaves:				
15	NP(+) ⁸	0.344	36 (36)	1:1
	NP(+) ⁹	0.327	2 (20)	1:10
	NP(+) ¹⁰	0.406	34 (33)	1:1
	NP(+) ¹¹	0.156	5 (20)	1:4
	NP(+) ¹²	0.133	9 (57)	1:6
	NP(+) ¹³	0.144	2 (7)	1:4
	NP(+) ¹⁴	0.040	0 (19)	0:19
20	NP(+) ¹⁶	0.053	0 (10)	0:10
5-6 leaves:				
25	NP(+) ²⁰	0.487	203 (117)	2:1
	NP(+) ²¹	0.042	0 (117)	0:117
	NP(+) ²²	0.142	0 (208)	0:208
	NP(+) ²³	0.317	223 (208)	1:1
	NP(+) ²⁴	0.051	0 (35)	0:35
	NP(+) ²⁵	0.286	13 (35)	1:3
	NP(+) ²⁶	0.037	0 (22)	0:22
30	NP(+) ²⁷	0.425	305 (22)	14:1

^aproduction of the NP in transgenic plants was assayed by double antibody sandwich enzyme-linked immunosorbent assay (DAS-ELISA); concentration of antibodies against-virus for coating: 1 µg/ml; dilution of conjugate to the NP of TSWV-BL: 1:250; results taken 150 min. after adding substrate; readings at 405 nm.

^blocal lesions that developed on inoculated leaves were counted 7 days after inoculation. Data represent the average of three inoculated leaves. Data in parentheses are the number of lesions produced from control NP(-) plants inoculated with the same inoculum.

^cthe ratio of local lesions that developed on NP(+) plants transformed with pBIN19-

NP⁺ versus local lesions that developed on the control NP(-) plant when inoculated with the same inoculum.

Of the 23 NP(+) clones, 10 produced high levels of NP, 5 accumulated intermediate levels of NP, and the remaining 8 produced low levels of NP. The size of the NP expressed in transgenic plants was analyzed using Western blot. Many polypeptides from tobacco extracts were reactive to the antibodies against the whole ^{virus} ~~virus~~ even though the antibodies were pre-absorbed with extracts from healthy tobacco plants. Of those, only one band was unique to the pattern of polypeptides from tobacco plants transformed with the NP gene. This polypeptide was estimated to be around 29 kDa, which is near the expected size of the native NP. No antibody reactive-protein band of similar size was found in extracts from transgenic plants containing the vector pBIN19.

Inoculation of tobacco leaves with TSWV-BL isolate could result in either systemic infection or necrotic local lesions, depending upon weather conditions and physiological stages of plants. When R₀ plants were tested with TSWV-BL for viral resistance, TSWV-BL induced typical necrotic lesions on the inoculated leaves of control NP(-) plants 6-8 days after inoculation. However, transgenic NP(+) plants showed a spectrum of resistance to the virus when compared to control NP(-) plants. Eleven of the 23 NP(+) plants did not develop any local lesion or the number of lesions that developed was at least 20-fold less than that on the corresponding inoculated NP(-) plants. Three NP(+) plants had intermediate reactions (5- to 19-fold less lesions than controls) while the remaining 9 plants had low or no resistance. None of the inoculated NP(+) or NP(-) plants showed systemic infection.

^{Symptomless} ~~symptomless~~ R₀ plants were monitored until the end of their life cycle, and no symptom was observed throughout their life cycles. The inoculated leaves of the symptomless NP(+) plants were checked for the presence of the virus on the leaves of *C. quinoa* plants. No virus was recovered from TSWV-BL-challenged leaves of highly resistant NP(+) plants, suggesting that the virus ^{could} ~~could~~ not replicate or spread in these NP(+) plants.

Leaf discs from selected R₀ plants were subcloned, and the regenerated plantlets were challenged by the virus. All subcloned R₀ plants displayed levels of resistance similar to their corresponding original R₀ plants.

5 Since TSWV is widespread and many biologically distant strains exist, the effectiveness of the transgenic plants to resist infections by different TSWV isolates were also tested. Five TSWV isolates were chosen in this study to challenge R₁ plants germinated on kanamycin-containing medium: TSWV-BL, Arkansas, 10W pakchoy, Begonia and
10 Brazil. The first three isolates were reactive to the antibodies against the whole virion and the NP of TSWV-BL (the common TSWV "L" serogroup) (see figure 5). Begonia isolate reacted strongly to the antibodies against the NP of TSWV-I (the "I" serogroup) but not to those raised against the TSWV-BL NP, and therefore belonged to the "I" serogroup. No detectable reaction of Brazil isolate was found to the
15 antibodies against either the NP of the TSWV-BL or the TSWV-I serogroup, and it was weakly reactive to the antibodies against the whole ^{virion} ~~NP~~ of TSWV-BL. Moreover, this isolate caused systemic mottle and crinkle on the leaves of infected tobacco and *N. benthamiana*,
20 but did not infect squash or cucumbers indicating that it is a distinct isolate from the cucurbit isolate. These results indicate that this isolate may be considered to be a third serogroup.

Seedlings derived from seven R₀ lines were germinated on kanamycin medium and inoculated with the above TSWV isolates.
25 Infectivity data were recorded daily starting seven days after inoculation. Plants inoculated with TSWV-BL, Arkansas, 10W pakchoy or Brazil isolates were rated susceptible if virus symptoms were observed on uninoculated leaves. Plants inoculated with the Begonia isolate were rated susceptible if local lesions were observed on
30 inoculated leaves, since this isolate does not cause systemic infection in tobacco. All inoculated control NP(-) R₁ plants were susceptible to infection by these five isolates. They were systemically infected 12 days after inoculation except that transgenic R₁ plants inoculated with Begonia produced only local lesions on the inoculated leaves. However,
35 almost all NP(+) R₁ plants were highly resistant to the homologous

Isolate TSWV-BL, while much lower percentages of NP(+) R₁ plants were resistant to heterologous isolates Arkansas, 10W pakchoy and Begonia. On the other hand, all NP(+) R₁ plants from the seven transgenic lines were susceptible to the Brazil isolate, even though a slight delay (1 to 2 days) in symptom expression was observed in some of the high NP-expressing NP(+) R₁ plants from line NP(+)₄.

Resistant R₁ plants remained symptomless throughout their life cycles. The inoculated leaves of seventeen ^{symptomless} ~~symptomless~~ NP(+) plants were checked for the presence of the virus by back inoculation on leaves of *Chenopodium quinoa* plants. No virus was recovered from the inoculated leaves of symptomless NP(+) plants, suggesting that the virus could not replicate or spread in these NP(+) plants.

The relationship between the level of NP accumulation in transgenic plants and the degree of resistance to heterologous TSWV isolates was also studied. Analysis of the data described above suggested that R₁ plants derived from R₀ lines with low levels of NP offered the best resistance to the heterologous isolates of the "L" serogroup (Arkansas and 10W pakchoy) while R₁ from a R₀ line with high level of NP were resistant to the Begonia isolate, which belongs to the "I" serogroup. For example, an average 76% of inoculated R₁ plants from low NP expressing lines NP(+) 2, 14, and 21 were resistant to infections by the Arkansas and 10W pakchoy isolates, while resistance to these isolates was observed in only 11% of similarly inoculated plants from high NP expressing lines NP(+)₄, 9, and 23. On the other hand, the Begonia isolate infected 79% of R₁ plants from the low NP expressing line NP(+)₂, 14, and 21 but only 19% from high NP expressing line NP(+)₄.

Therefore, it was concluded that the transgenic R₁ plants expressing low levels of the NP gene were highly resistant to infection with the isolate 10W pakchoy (the "L" serogroup), but not to Begonia isolate (the "I" serogroup). In contrast, the highly NP-expressing R₁ plants were very resistant to infection by Begonia isolate but not to infection by the isolate from 10W pakchoy.

Thus, it was of interest to accurately quantitate the relation of NP expression in individual plants with resistance to the heterologous

isolates. In a number of inoculation experiments reported herein, leaf samples of transgenic plants were taken before inoculating with the Arkansas and 10W pakchoy isolates. Samples were also taken from non-inoculated leaves of plants inoculated with the Begonia isolate after observations of the apparent relation between NP expression levels and resistance were made. The latter method of sampling could be done without interference from infection by the Begonia isolate because this isolate does not cause systemic infection in tobacco nor reacts with antibodies to the TSWV-BL NP. All samples were assayed for relative NP levels by DAS-ELISA using antibodies raised to isolated NP of TSWV-BL. Figures 5 and 6 show the relation between NP levels in transgenic R₁ plants (irrespective of the R₀ lines they came from) and their resistance to the Arkansas and 10W pakchoy isolates or to the Begonia isolate. Nearly all transgenic R₁ plants with very low or undetectable ELISA reactions (0-0.05 OD_{405nm}) were resistant to infections by the Arkansas and 10W pakchoy isolates (the "L" serogroup) but susceptible to the Begonia isolate (the "I" serogroup). In contrast, almost all R₁ plants that gave high ELISA reactions (0.4-1.0 OD_{405nm}) were resistant to the Begonia isolate but susceptible to the Arkansas and 10W pakchoy isolates.

The double-stranded (ds) RNA was isolated from the *N. benthamiana* plants infected with TSWV-B using a combination of methods [See Acta Horticulturae 186:51 (1986), and Can. Plant Dis Surv 68:93(1988)] which have been successfully used for isolation of dsRNA from tissue infected with grapevine leafroll virus. The dsRNA was chosen for the cDNA synthesis since isolation of the virus particle from this isolate has not been possible [see Plant Disease 74:154 (1990)]. In order to make a cDNA library specific to the S RNA of TSWV-B, the double stranded S RNA was gel-purified, denatured by methyl-mercury treatment, and subjected to cDNA synthesis procedure provided by Promega using random primers. The synthesized cDNA fragments were cloned via an EcoRI adaptor into the EcoRI digested λ ZAPII (Stratagene), and positive clones were identified by colony hybridization using the cDNA probes prepared by reverse transcription of gel-purified S RNA. Dozens of positive clones were analyzed on

agarose gels and only three overlapping clones containing the largest inserts (L1, L22 and L30) were selected (see figure 3), covering nearly entire TSWV-B S RNA.

5 The nucleotide sequences of the inserts in clones L1, L22 and L30 were determined from both DNA strands, first by the universal and reverse primers and then by the internal primers designed for sequencing the S RNA of TSWV-B. Sequencing was done using the Sanger dideoxyribonucleotide method, T7 polymerase (U.S. Biochemicals, Sequenase TM), and the double-stranded sequencing
10 procedure described by Slemieniak [see Analyt. Biochem. 192:441 (1991)]. The sequence analyses of these clones revealed inserts of 1.994 kb, 2.368 kb and 1.576 kb, respectively, and these sequences represented 93% of the S RNA genome (see figure 3). The assembled sequence was analyzed by comparisons with sequences of TSWV
15 isolates CONH1, L3, I, and BL using computer programs available from the Genetics Computer Group (GCG, Madison, WI).

Computer analysis showed that the assembled sequence of 2.842 kb covered the complete 52 K nonstructural protein gene, the complete intergenic region (629 bp), and 737 bp of the NP gene (only 39 N-
20 terminal nucleotides of the N gene were not represented). In order to obtain this missing region of the N gene, a primer
(SEQ ID NO. 10) TTCTGGTCTTCTTCAAAC^(SEQ ID NO. 10)TCA, identical to a sequence 62 nucleotides from the initiation codon of the N gene, was end-labeled with polynucleotide kinase to screen the cDNA library described above. Five
25 putative clones were obtained. Sequence analysis of the five clones showed that only clones S6 and S7 contain these 39 missing nucleotides of the N gene. The latter clone also included the extreme 3' end of the S RNA.

The 5' extreme end of the S RNA was obtained using the 5' RACE
30 System (GIBCO). Both ssRNA of TSWV-B and total RNAs isolated from tobacco plants infected with TSWV-B were used to synthesize first strand cDNA with an oligonucleotide (5'-CTGTAGCCATGAGCAAAG)
(SEQ ID NO. 11) complementary to the nucleotide positions 746-763 of the TSWV-B S RNA. The 3'-end of the first strand cDNA was tailed with dCTP using
35 terminal deoxynucleotidyl transferase. Tailed cDNA was then amplified

by PCR using an ^{anchor} primer that anneals to the homopolymeric tail, and an oligonucleotide (5'-TTATATCTTCTTCTTGGA) ^(SEQ ID NO. 17) that anneals to the nucleotide positions 512-529 of the TSWV-B S RNA. The PCR-
^{amplified fragment} amplified fragment was gel-purified and directly cloned into the T-
 5 vector pT7Blue (Novagen) for sequence analysis. Eight independent clones were sequenced with an oligomer primer (5'-
 GTTCTGAGATTGCTAGT) ^(SEQ ID NO. 16) close to the 5' region of the S RNA (nucleotide positions 40-57 of the TSWV-B S RNA). Six of the resulting clones contained the 5' extreme end of the S RNA and the 5'-terminal
 10 nucleotide sequence from these clones was identical. Thus, the complete nucleotide sequence of the TSWB-B S RNA is 3049 nucleotides in length.

Thus these two clones together with the three clones previously
 15 sequenced (L1, L22, L30, S6 and S7) covered a total of 3032 nucleotides depicted above. Comparisons with the terminal sequences of TSWV-CPNH1 and TSWV-I revealed that although the extreme 5' end of 18 nucleotides was not represented in the assembled sequence, the extreme 3'-terminus of the TSWV-B S RNA is identical to the extreme 3' end of the TSWV-I S RNA and is only one out of fifteen nucleotides
 20 different from the extreme 3' end of TSWV-CPNH1. The conservation of the terminal sequence among TSWV isolates is consistent with observations of the other members of *Bunyaviridae* genera, and supports the hypothesis that the terminal sequences might form stable base-paired structure, which could be involved in its replication and
 25 encapsulation.

The complete nucleotide sequence of the S RNA genome of TSWV-B (the Brazilian isolate discussed above) according to the present
 invention is: ^(SEQ ID NO. 14)

	AGAGCAATTG GGICATTTT TATTCTAAAT CGAACCTCAA CTAGCAAATC	50
30	TCAGAACTGT AATAAGCACA AGAGCACAAG AGCACAATG TCATCAGGTG	100
	<u>TTTATGAATC GATCATTGAG ACAAGGCTT CAGTTTGGGG ATCGACAGCA</u>	150
	<u>TCTGGTAAGT CCATCGTGGG TCTTACTGG ATTTATGAGT TTCCAACCTG</u>	200
	<u>TTCTOCACCTG GTTCAAACCTC AGTTGTACTC TGATTGAGG AGCAAAAGTA</u>	250
	<u>GCTTGGGCTA CACTTCAAAA ATTGGTGATA TTCTGCTGT AGAGGAGGAA</u>	300
35	<u>ATTTTATCTC AGAACGTTC TATCCAGTG TTGATGATA TTGATTTCAG</u>	350

	CATCAATATC AATGATTCTT TCTTGGCAAT TTCTGTTTGT TOCAACACAG	400
	TTAACACCAA TGGAGTGAAG CATCAGGGTC ATCTTAAAGT TCTTTCTCTT	450
	GOCCAATTGC ATGCTTTGA ACGTGTGATG AGCAGGTGAG AGATTGCTAG	500
	CAGATTGCGG GTGCAAGAAG AAGATATAAT TOCTGATGAC AAATATATAT	550
5	GTGGTGGTAA CAAGGGATGT GTCTGCTGTG TCAAGAACA TACTTACAAA	600
	GTGAAATGA GGCACAATGA GGGTTTAGGC AAAGTGAATG TCTTTCTGTC	650
	TAACAGAAAT GTTCATGAGT GGCTGTATAG TTCAAAACCA AATTTCBAAC	700
	AGATGAAAG TAATAACAGA ACTGTAAATT CTCTTGCAGT CAAATCTTTG	750
	CTCATGGCTA CAGAAAACAA CATTATGCGT AACTCTCAAG CTTTGTGTAA	800
10	AGCTTGTACT GATTCTCATT TTAAGTTGAG CCTTGGGCTG AGAATTGCAA	850
	AAGTTTGTAA GCAAATAGGC ATACAGAAGC TCTTCAAGTT TGCAGGAGAC	900
	GAAACGGTAA AAAGTTTGTAA TTGTCTTATT GCATGCATCC CAAATCACAA	950
	CAGTGTGGAA ACAGCTTTAA ATGTACACGT TATATGTAGA CATCAGCTTC	1000
	CAATGCGTAA GTGCAAGCT CCGTTTGAAT TATCAATGAT TTCTCTCGAT	1050
15	CTGAAAGAGC CTTACAAGAC TGTGCAATGAT CCTTCATATC CTCAAAGGAT	1100
	TGTTTATGCT TTGGTTGAGA CTCACACTTC CTTTGCACAA GTTCTCTGCA	1150
	ACAAGCTGCA AGAAGATGTG ATCATATATA CTATAAACAG CCTGAACTA	1200
	AOCOCAGCTA AGCTGGATCT AGGTGAAAGA ACCTTGAAC TACAGTGAAGA	1250
	TGCTTCGAAG AAGAAGTATT TTCTTTTCAA AACACTGGA TGCTTGGCAG	1300
20	TAAATGTGCA GACTATGTCT TATTTGGATA GCATCCAGAT TOCTTCATGG	1350
	AAGATAGACT TTGCCAGAGG AGAGATCAGA ATCTCCCTC AATCTACTCC	1400
	TATTGCAAGA TCTTTGCTCA AGCTGGATTT GAGCAAGATC AAGGAAAGA	1450
	AGTCTTIGAC TTGGGAAACA TCCAGCTATG ATCTAGAATA AAAGTGGCTC	1500
	ATACTACTCT AAGTAGTATT TGTCAACTTG CTTATCCTTT ATGTTGTTTA	1550
25	TTTCTTTTAA ATCTAAAGTA AGTTAGATT CAGTAGTTTA GTATGCTATA	1600
	GCATTATTAC AAAAAATACA AAAAAATACA AAAAAATACA AAAAAATATA	1650
	AAAAOCCAAA AAGATCCCAA AAGGGACGAT TTGGTTGATT TACTCTGTTT	1700
	TAGGCTTATC TAAGCTGCTT TTGTTTGGC AAAATAACAT TGTAACATGC	1750
	AATAACTGGA ATTTAAAGTC CTAAAAGAAG TTTCAAAGGA CAGCTTAGCC	1800
30	AAAATTGGTT TTGTTTGTG TTTTGTGTG TTTTGTGTG TTGTTTATT	1850
	TTTATTTTAA GTTATTTTT TTGTTTGTG ATTTTATTT TTATTTTATT	1900
	TTCTTTTATT TTATTTATAT ATATATCAA CACAATCCAC ACAAATAATT	1950
	TTAATTTCAA ACATTCTACT GATTTAACAC ACTTAGCCTG ACTTTATCAC	2000
	ACTTAACACG CTTAGTTAGG CTTTAACACA CTGAAGTGA TTAACACACA	2050
35	CTTAGTATTA TGCACTCTT AATTAAACA CTTAATAAT ATGCATCTCT	2100

001253-10300

```

GAATCAGCCT TAAAGAAGCT TTTATGCAAC ACCAGCAATC TTGGCCTCTT 2150
TCTTAACCTCC AACCATTTCA TAGAATTTGT CAAGATTATC ACTGTAATAG 2200
TOCATAGCAA TGCTTCOCTT AGCATTGGGA TTGCAAGAAC TAAGTATCTT 2250
GGCATATTCT TTCCCTTTGT TTATCTGTGC ATCATOCATT GTAAATCCTT 2300
5 TGCTTTTAAAG CACTGTGCAA AOCTTCCOCA GAGCTTCCTT AGTGTGTGAC 2350
TTAGTTGGTT CAATCCCTAA CTCCCTGTAC TTTGCATCTT GATATATGGC 2400
AAGAACAACA CTGATCATCT CGAAGCTGTC AACAGAAGCA ATGAGAGGGA 2450
TACTAACCTCC AAGCATTATA GCAAGCTCA CAGATTTTGC ATCTGCCAGA 2500
GGCAGCCCGT AAGCTTGGAC CAAAGGGTGG GAGGCAATTT TTGCTTTGAT 2550
10 AATAGCAAGA TTCTCATTGT TTGCAGCTTC TTCTATGAGC TTCACTCTTA 2600
TCATGCTATC AAGCCTCCTG AAAGTCATAT CCTTAGCTCC AACTCTTTCA 2650
GAATTTTCT TTATCGTGAC CTTACCAAAA GTAAAATCAC TTGGTTTCAC 2700
AACTTTCATA ATGCCTTGGC GATTCTTCAA GAAAGTCAAA CATGAAGTGA 2750
TACTCATTTT CTTAATCAGG TCAAGATTTT CCTGACAGAA AGTCTTAAAG 2800
15 TTGAATGCGA CCTGGTTCIG GTCTTCTTCA AACTCAACAT CTGCAGATTG 2850
AGTTAAAAGA GAGACAATGT TTTCTTTTGT GAGCTTGACC TTAGACATGG 2900
TGGCAGTTTA GATCTAGACC TTTCTCGAGA GATAAGATT C AAGGTGAGAA 2950
AGTGAACAC TGTAGACCGC GGTCGTACT TATCCGTGA ATGTGATGAT 3000
TTGTATTGCT GAGTATTAGG TTTTGAATA AAATTGACAC AATTGCTCT 3049

```

2.0 The deduced amino acid sequences of the nonstructural (single underlined above) and nucleocapsid proteins according to the present invention are: (SEQ ID NO. 12)

SUB A 6

Met	Ser	Ser	Gly	Val	Tyr	Glu	Ser	Ile	Ile	Gln	Thr	Lys	Ala	Ser	5	10	15
Val	Trp	Gly	Ser	Thr	Ala	Ser	Gly	Lys	Ser	Ile	Val	Asp	Ser	Tyr	20	25	30
Trp	Ile	Tyr	Glu	Phe	Pro	Thr	Gly	Ser	Pro	Leu	Val	Gln	Thr	Gln	35	40	45
Leu	Tyr	Ser	Asp	Ser	Arg	Ser	Lys	Ser	Ser	Phe	Gly	Tyr	Thr	Ser	50	55	60
Lys	Ile	Gly	Asp	Ile	Pro	Ala	Val	Glu	Glu	Glu	Ile	Leu	Ser	Gln	65	70	75
Asn	Val	His	Ile	Pro	Val	Phe	Asp	Asp	Ile	Asp	Phe	Ser	Ile	Asn	80	85	90
Ile	Asn	Asp	Ser	Phe	Leu	Ala	Ile	Ser	Val	Cys	Ser	Asn	Thr	Val	95	100	105
Asn	Thr	Asn	Gly	Val	Lys	His	Gln	Gly	His	Leu	Lys	Val	Leu	Ser	110	115	120

SUB A 6

00126788-10300

	Leu	Ala	Gln	Leu	His	Pro	Phe	Glu	Pro	Val	Met	Ser	Arg	Ser	Glu
					125					130					135
	Ile	Ala	Ser	Arg	Phe	Arg	Leu	Gln	Glu	Glu	Asp	Ile	Ile	Pro	Asp
					140					145					150
5	Asp	Lys	Tyr	Ile	Ser	Ala	Ala	Asn	Lys	Gly	Ser	Leu	Ser	Cys	Val
					155					160					165
	Lys	Glu	His	Thr	Tyr	Lys	Val	Glu	Met	Ser	His	Asn	Gln	Ala	Leu
					170					175					180
10	Gly	Lys	Val	Asn	Val	Leu	Ser	Pro	Asn	Arg	Asn	Val	His	Glu	Trp
					185					190					195
	Leu	Tyr	Ser	Phe	Lys	Pro	Asn	Glu	Asn	Gln	Ile	Glu	Ser	Asn	Asn
					200					205					210
	<u>Arg</u>	<u>Thr</u>	<u>Val</u>	<u>Asn</u>	<u>Ser</u>	<u>Leu</u>	<u>Ala</u>	<u>Val</u>	<u>Lys</u>	<u>Ser</u>	<u>Leu</u>	<u>Leu</u>	<u>Met</u>	<u>Ala</u>	<u>Thr</u>
					215					220					225
15	Glu	Asn	Asn	Ile	Met	Pro	Asn	Ser	Gln	Ala	Phe	Val	Lys	Ala	Ser
					230					235					240
	Thr	Asp	Ser	His	Phe	Lys	Leu	Ser	Leu	Gln	Leu	Arg	Ile	Pro	Lys
					245					250					255
20	Val	Leu	Lys	Gln	Ile	Ala	Ile	Gln	Lys	Leu	Phe	Lys	Phe	Ala	Gly
					260					265					270
	Asp	Glu	Thr	Gly	Lys	Ser	Phe	Tyr	Leu	Ser	Ile	Ala	Cys	Ile	Pro
					275					280					285
	Asn	His	Asn	Ser	Val	Glu	Thr	Ala	Leu	<u>Asn</u>	<u>Val</u>	<u>Thr</u>	<u>Val</u>	Ile	Cys
					290					295					300
25	Arg	His	Gln	Leu	Pro	Ile	Pro	Lys	Ser	Lys	Ala	Pro	Phe	Glu	Leu
					305					310					315
	Ser	Met	Ile	Phe	Ser	Asp	Leu	Lys	Glu	Pro	Tyr	Asn	Thr	Val	His
					320					325					330
30	Asp	Pro	Ser	Tyr	Pro	Gln	Arg	Ile	Val	His	Ala	Leu	Leu	Glu	Thr
					335					340					345
	His	Thr	Ser	Phe	Ala	Gln	Val	Leu	Cys	Asn	Lys	Leu	Gln	Glu	Asp
					350					355					360
	Val	Ile	Ile	Tyr	Thr	Ile	Asn	Ser	Pro	Glu	Leu	Thr	Pro	Ala	Lys
					365					370					375
35	Leu	Asp	Leu	Gly	Glu	Arg	Thr	Leu	<u>Asn</u>	<u>Tyr</u>	<u>Ser</u>	<u>Glu</u>	<u>Asp</u>	<u>Ala</u>	<u>Ser</u>
					380					385					390
	Lys	Lys	Lys	Tyr	Phe	Leu	Ser	Lys	Thr	Leu	Glu	Cys	Leu	Pro	Val
					395					400					405
40	Asn	Val	Gln	Thr	Met	Ser	Tyr	Leu	Asp	Ser	Ile	Gln	Ile	Pro	Ser
					410					415					420
	Trp	Lys	Ile	Asp	Phe	Ala	Arg	Gly	Glu	Ile	Arg	Ile	Ser	Pro	Gln
					425					430					435
	Ser	Thr	Pro	Ile	Ala	Arg	Ser	Leu	Leu	Lys	Leu	Asp	Leu	Ser	Lys
					440					445					450

SUB A
cont

Ile Lys Glu Lys Lys Ser Leu Thr Trp Glu Thr Ser Ser Tyr Asp
455 460 465

Leu Glu;
and (SEQ ID NO. 13)

5 Met Ser Lys Val Lys Leu Thr Lys Glu Asn Ile Val Ser Leu Leu
5 10 15
Thr Gln Ser Ala Asp Val Glu Phe Glu Glu Asp Gln Asn Gln Val
20 25 30
Ala Phe Asn Phe Lys Thr Phe Cys Gln Glu Asn Leu Asp Leu Ile
35 40 45
10 Lys Lys Met Ser Ile Thr Ser Cys Leu Thr Phe Leu Lys Asn Arg
50 55 60
Gln Gly Ile Met Lys Val Val Asn Gln Ser Asp Phe Thr Phe Gly
65 70 75
15 Lys Val Thr Ile Lys Lys Asn Ser Glu Arg Val Gly Ala Lys Asp
80 85 90
Met Thr Phe Arg Arg Leu Asp Ser Met Ile Arg Val Lys Leu Ile
95 100 105
Glu Glu Thr Ala Asn Asn Glu Asn Leu Ala Ile Ile Lys Ala Lys
110 115 120
20 Ile Ala Ser His Pro Leu Val Gln Ala Tyr Gly Leu Pro Leu Ala
125 130 135
Asp Ala Lys Ser Val Arg Leu Ala Ile Met Leu Gly Gly Ser Ile
140 145 150
25 Pro Leu Ile Ala Ser Val Asp Ser Phe Glu Met Ile Ser Val Val
155 160 165
Leu Ala Ile Tyr Gln Asp Ala Lys Tyr Lys Glu Leu Gly Ile Glu
170 175 180
Pro Thr Lys Tyr Asn Thr Lys Glu Ala Leu Gly Lys Val Cys Thr
185 190 195
30 Val Leu Lys Ser Lys Gly Phe Thr Met Asp Asp Ala Gln Ile Asn
200 205 210
Lys Gly Lys Glu Tyr Ala Lys Ile Leu Ser Ser Cys Asn Pro Asn
215 220 225
35 Ala Lys Gly Ser Ile Ala Met Asp Tyr Tyr Ser Asp Asn Leu Asp
230 235 240
Lys Phe Tyr Glu Met Phe Gly Val Lys Lys Glu Ala Lys Ile Ala
245 250 255
Gly Val Ala

40 As the nucleocapsid protein gene depicted above is on the viral
(SEQ ID NO. 15)
complementary strand, the nucleocapsid protein gene of TSWV-B is:

ATG TCT AAG GTC AAG CTC ACA AAA GAA AAC ATT GTC TCT CTT TTA 45
ACT CAA TCT GCA GAT GTT GAG TTT GAA GAA GAC CAG AAC CAG GTC 90

SUB A

gene Comparisons ^a	Overall 53 K protein gene			Intergenic	29 K protein	
	nt	nt	aa	nt	nt	aa
5 B/CPNH1	76.4 ^b	80.0	86.1(78.3) ^c	72.4	77.5	91.5(79.1)
B/L3	75.8	79.0	89.0(82.0)	76.4	78.0	91.1(79.9)
B/BL	76.3	-	-	72.8	77.6	90.3(79.5)
B/I	63.0	-	-	-	63.1	69.7(55.3)
CPNH1/L3	94.8	95.6	92.0(89.4)	89.2	96.8	99.6(98.5)
CPNH1/BL	96.4	-	-	95.9	97.2	98.8(96.9)
10 CPNH1/I	62.7	-	-	-	60.8	69.5(55.1)
L3/BL	95.1	-	-	92.6	97.3	99.2(98.5)
L3/I	60.9	-	-	-	60.9	69.5(55.1)
I/BL	61.7	-	-	-	60.9	68.8(53.9)

15 ^a The partial or complete S RNA sequences of isolates TSWV-CPNH1 (2.916 kb), TSWV-L3 (2.837 kb), TSWV-BL (2.037 kb) and TSWV-I (1.144 kb) were used for comparisons with the S RNA sequence of the TSWV-B (3.049 kb).

^b Percent similarities were calculated by Comparison of their nucleotide or predicted amino acid sequence using the program BESTFIT of the GCG Sequence analysis software package.

20 ^c Percent identity is in parenthesis.

As depicted, the greatest nucleotide sequence similarity (75.8%-76.4%) was shown with the L-type isolates (CHNH1, L3 and BL). To the lesser extent, there was nucleotide sequence similarity (63%) between the TSWV-B S RNA and the S RNA of TSWV-I assigned to I serogroup.

25 For comparison, the sequenced S RNA regions of the L-type isolates (CHPN1, L3 and BL) shared 94.8%-96.4% nucleotide sequence similarities.

30 The open reading frame of 777 nucleotides encodes the N protein of 258 amino acids with a predicted molecular weight of 28700 Da. ^{The} sequence comparisons of the N open reading frame from TSWV isolates revealed that nucleotide sequences of the N genes from the isolates CPNH1, L3 and BL ^{differ} ~~differs~~ from TSWV-B by a considerably larger amount (22%-22.5%) than they differ from each other (2.7%-3.2%). Consistent to the results of the immunological analysis, the N amino acid

35 sequences among CPNH1, L3 and BL isolates are more closely related to each other (98.8%-99.6% similarities or 96.9%-98.5% identities) than to

the TSWV-B (90.3%-91.5% similarities or 79.1%-79.9% identities). Much lower homology was observed to TSWV-I at both nucleotide (63.1%) and amino acid (69.7% similarity or 55.3% identity) levels. Except for the N open reading frame of TSWV-I that encodes 262 amino acids, the N open reading frames of the other isolates code for the 258 amino acids. Computer analysis suggested that the extra residues of TSWV-I N open reading frame resulted from the amino acid sequence insertions (residues 82 through 84 and residue 116). One potential N-glycosylation site is found at residue 68.

The second open reading frame of 1404 nucleotides encodes the nonstructural protein of 467 amino acids with a predicted molecular weight of 52566 Da. Comparisons with homologous open reading frames of TSWV-CPNH1 and TSWV-L3 showed 80% and 79% similarities at the nucleotide level, and 86.1% (or 78.3% identity) and 89% (or 82.0% identity) similarities at the amino acid level. This open reading frame contains four potential glycosylation sites, which are located in the ~~exactly same positions~~ ^{exact same position} as those of TSWV-CPNH1 and TSWV-L3.

The intergenic region of the TSWV-B S RNA was, due to several insertions, 126 and 41 nucleotide longer than the counterparts of TSWV-CPNH1 and TSWV-L3, respectively. The sequence analysis by the program FOLD indicated the intergenic region can form very complex and stable hairpin structure by internally base-pairing U-rich stretches with A-rich stretches of the intergenic region, which had similar stability to those produced from TSWV-CPNH1 and TSWV-L3 as indicated by minimum free energy values. This internal base-paired structure may act as a transcription termination signal.

The results tabulated above also revealed that the N protein of TSWV-B is subject to a higher degree of selective pressure than the 52 K protein; the similarities among the amino acid sequences of the 52 K protein are lower than that found for the amino acid sequence of the NPs. Nucleotide sequence divergence is highest among the intergenic regions, which indicates that this region is subject to less selective pressure than either genetic region.

The evolutionary relationships among the TSWV-B and other four TSWV isolates were analyzed and depicted in figure 4 in which the

evolutionary tree organization is consistent with the relatedness of serological data collected for these TSWV isolates. Thus, the TSWV-B, according to the present invention, is more closely related to the L-type isolates than to the I-type isolate TSWV-I, but is much less similar to the L-type isolates than the L-type isolates are to each other.

Despite a slight delay of symptom expression, transgenic plants did not show resistance to the Brazil isolate of TSWV; Serological results show that this isolate is distinct from the "L" and "I" type isolates, and biologically different from the cucurbit isolate. The Brazil isolate may thus belong to still another serogroup of TSWV. In any ~~event~~^{event}, infectivity results show that it is unlikely that a single NP gene will provide resistance to all isolates in the ~~Tospovirus~~^{Tospovirus} genus.

Transgenic plants according to the present ~~invention~~^{invention} that gave low or undetectable ELISA reactions (0-0.05 OD_{405nm}) were resistant to infection by the heterologous isolates (Arkansas and 10W pakchoy) of the "L" serogroup, whereas no protection against these isolates was found in plants accumulating high levels of the NP. Compared to the ELISA readings of control NP(-) plants (0.05 OD_{405nm}), these transgenic plants may produce little, if any, TSWV-BL NP. Similar results have been observed in transgenic plants, in which the CP accumulation was not detected; these were highly resistant to virus infection. The mechanism underlying this phenomenon is presently unknown. It is likely that this type of resistance might be attributed to interference of CP RNA molecules produced in transgenic plants with viral replication, presumably by hybridizing to minus-sense replicating RNA of the attacking virus, binding to essential host factors (e.g., replicase) or interfering with virion assembly.

It should be noted, however, that the resistance to the homologous TSWV-BL isolate is apparently independent of the expression levels of the NP gene. Although the relative NP levels of the individual R₁ plants inoculated with TSWV-BL were not measured, it is reasonable to assume that the NP produced in these inoculated R₁ plants (a total of 145 plants tested) ranged from undetectable to high.

In contrast to the case for protection against the heterologous isolates of the "L" serogroup, protection against the Begonia isolate of the TSWV-I serogroup was found in the high NP-expressing R₁ plants. Comparison of NP nucleotide sequence of the "L" serogroup with that of the "I" serogroup revealed 62% and 67% identity at the nucleotide and amino acid levels, respectively. The difference of NP genes of the two serogroups might be so great that the NP (the "L" serogroup) produced in transgenic plants acted as a dysfunctional protein on the attacking Begonia isolate of the "I" serogroup. Incorporation of this "defective" coat protein into virions may ^{generate} defective virus which inhibit virus movement or its further replication. This type of interaction is expected to require high levels of the NP for the protection.

Alternatively, resistance to the Begonia isolate may also involve interference of NP transcripts produced in R₁ plants with viral replication. If this is true, more NP transcripts (due to the heterologous nature of two NP gene) may be required to inhibit replication of heterologous virus.

Although there are no obvious explanations for the results showing the relation of NP levels in individual R₁ plants to resistance to the heterologous isolates of the "L" and "I" serogroups, it is believed these are definite trends since the data were derived from a large number (190) of plants. Thus, it is believed that a measurement of CP or NP levels in individual plants may provide a more accurate way to relate NP or CP levels to resistance. By this form of data analysis, the results show that the resistance was more closely related to NP levels in each test plant than to the NP level of the R₀ line from which they were derived. For TSWV-BL Np gene in tobacco, at least, it appears that integration sites of the NP gene in plant chromosome may not be important for viral resistance.

Studies have also been conducted to determine the reaction of transgenic R₁ and R₂ tomatoes containing the nucleocapsid protein gene of TSWV-BL according to the present invention to the following isolates: Brazil (a distantly related virus), T91 (a closely related virus) and BL (a homologous isolate). In these studies, transgenic tomatoes (*L. esculentum*) were produced by *A. tumefaciens*-mediated gene transfer

of the nucleocapsid protein (N) gene of the lettuce isolate of tomato spotted wilt virus BL into germinated cotyledons using modifications of published procedures [see Plant Cell Reports 5:81 (1986)]. The tomato line "Geneva 80" was selected for transformation because it contains the Tm-22 gene which imparts resistance to TMV, thus creating the possibility of producing a multiple virus-resistant line.

Transformants were selected on kanamycin media and rooted transgenic tomatoes were potted and transferred into the greenhouse. R₁ and R₂ tomato seedlings expressed the NPT II gene, suggesting multiple insertions of this gene in the plant genome. In contrast, only 18% of the seedlings produced detectable levels of the N protein.

Nine R₁ and three R₂ lines were tested for resistance to the following three *Tospovirus* described, specifically TSWV-BL, TSWV-T91, and TSWV-B. Infectivity was based upon visual inspection of test plants. In those cases where plants appeared healthy except for a few rust-colored rings or insect damage, extracts from these plants were inoculated to *N. benthamiana* to test for the presence of the virus. As depicted in the following table, nearly all control tomato plants exhibited typical symptoms consisting of plant stunting, leaf yellow mosaic and rugosity 3 to 4 weeks after inoculations with TSWV-BL, TSWV-T91 or TSWV-B. However, only 4% of the R₁ and R₂ transgenic plants became infected with TSWV-BL, 7% with TSWV-T91, and 45% with TSWV-B.

Viral resistance in transgenic R1 and R2 tomatoes expressing the nucleoprotein gene of the lettuce strain of tomato spotted wilt virus

		Inoculating Isolates ^a		
Plant Line		TSWV-BL	TSWV-T91	TSWV-B
5	R1 Plants:			
	T13-1	0/22	1/26	7/24
	T13-2	6/20	NT ^b	NT
	T13-3	2/42	0/20	12/18
	T13-4	0/25	NT	NT
10	T13-9	0/20	NT	NT
	T13-10	1/50	2/26	11/26
	T13-11	0/22	NT	NT
	T13-12	1/29	NT	NT
	T13-13	0/22	NT	NT
15	TOTAL	10/252	3/72	30/68
	R2 Plants:			
	T13-1-7	0/8	2/8	5/8
	T13-1-9	0/8	1/8	2/8
	T13-1-11	0/8	1/9	5/9
20	TOTAL	0/24	4/25	12/25
	CONTROLS	92/95	51/53	52/53

^a plants were inoculated at the one- to two-leaf stage with 5-, 10-, or 20-fold diluted leaf extract of *N. benthamiana*, H423 tobacco or tomato; the same plants were re-inoculated 7 days later and symptoms were recorded after another 14 days; the reaction is expressed as number of plants with symptoms/number of plants tested

^b not tested

Accordingly, the description above supports the finding that transgenic tomato plants that express the N gene of TSWV-BL show resistance to infection to TSWV-BL, to other TSWV isolates that are closely related to TSWV-BL, and to the more distantly related TSWV-B.

In further limited studies with an additional isolate, all transgenic plants were resistant to the 10W (pakchoy) isolate, whereas the controls were infected. These results show that transgenic tomatoes are better protected against closely related isolates than distantly related isolates. Unlike in transgenic tobacco and *N. benthamiana* expressing the TSWV-BL N gene, the level of N protein expression did not correlate with the observed protection in transgenic

tomatoes; 55% of the transgenic tomatoes were also resistant to a distantly related isolate of TSWV-B, which was not observed in transgenic tobacco and *N. benthamiana* plants. These discrepancies may reflect that tomato is inherently less susceptible to *Tospoviruses*.

5 In addition, studies were also conducted to determine virus distribution in a small number of plants at 5 and 7 weeks after inoculation. The distal halves from leaflets of all expanded leaves of each plant were ground and back-inoculated onto *N. benthamiana*. The results taken seven days after inoculation showed that virus cannot be
10 recovered from any leaf tissue of asymptomatic transgenic plants inoculated with either TSWV-BL, -T91, or -B, confirming the visual findings reported above. In transgenic plants showing symptoms, the virus is not distributed throughout the plant. For example, a transgenic plant which could not be conclusively rated visually contained the virus
15 in only two of the 8 leaves; the second leaves from the bottom and top of the plant. Conversely, virus^{was} present in all leaves of the infected control plant, ^{and was} ~~and is~~ absent in those of the healthy control plants.

Graft inoculations were attempted to test whether the resistant transgenic plants could become infected if virus is introduced into the
20 vascular system. R₁ and R₂ plants that had been inoculated at 1:5, 1:10 or 1:20 dilutions of TSWV-BL, -T91, or -B were grafted onto control plants infected with the same isolates and dilutions. The 34 transgenic plants were asymptomatic after 31 days, although the non-transgenic controls were infected. After 23 days, the top 46 cm of transgenic
25 plants had been trimmed away to induce new growth and more plant stress. Although the young, vigorously growing new shoots failed to show any symptoms on the 31st day post inoculation, 33%, 31% and 45% of TSWV-BL, -T91 and -B were showing leaf or stem symptoms, respectively at 45 days post inoculation. These results indicate that
30 some transgenic plants are tolerant, and others are immune to infection.

Thus, according to one aspect of the present invention, transgenic plants expressing the NP gene of the TSWV-BL isolate are highly
35 resistant to infections of both the homologous TSWV-BL isolate and heterologous isolates of the same serogroup (Arkansas and 10W

pakchoy). More significantly, the resistance is effective to Begonia isolate from other serogroups. In brief, the above clearly describes that transgenic tobacco plants expressing the nucleoprotein gene of TSWV-BL display resistance to both TSWV and INSV, and the protection appears to be mediated by the nucleoprotein against distantly related INSV and by the nucleoprotein gene ribonucleotide sequence against the homologous and closely related TSWV isolates. This is the first time broad spectrum resistance of the engineered plants to different isolates of TSWV has been shown.

While coat protein protection generally displays delay and/or reduction in infection and symptom expression, but no immunity, the present invention provided a significantly high percentage of transgenic plants which were symptom-free and free of the infective virus. Resistance of these plants under greenhouse conditions persisted throughout their life cycle, and more importantly was inherited to their progenies as shown above.

It was observed in the present invention that the transgenic plants producing little, if any, TWSV-BL NP were highly resistant to infection by the homologous isolate and other closely-related isolates within the same serogroup of TSWV, whereas no protection was found in those expressing high levels of the NP gene.

The biological diversity of TSWV is well documented and has been reported to overcome the genetic resistance in cultivated plants such as tomato. Thus, it is extremely important to develop transgenic plants that show ^{resistance} resistant to many strains of TSWV. The present invention indicates that one method to do so would be to utilize the viral NP gene to confer this resistance, and that this resistance would be to diverse TSWV isolates. Thus, the finding of the present invention that the expression of TSWV NP gene is capable of conferring high levels of resistance to various TSWV isolates has a great deal of commercial importance.

In another series of studies, Plasmid BIN19-N⁺ was constructed and transferred to *A. tumefaciens* strain LBA4404 in accordance with Example IV, and transferred to *Nicotiana benthamiana* in accordance

with Example V. The nucleocapsid genes of INSV-Beg and -LI were amplified with oligomer primers INSV-A

(5'-TAGTTATOTAGAAGCATGGACAAAGCAAAGATTACCAAGG) (SEQ. ID NO. 20) and INSV-B
(5'-TAGAGTGGATCGATGGTTATTTCAAATAATTTATAAAAAGCAO) (SEQ. ID NO. 21)

5 hybridizing to the 5'-coding and 3'-noncoding regions of the nucleocapsid gene of an INSV isolate, respectively. The amplified nucleocapsid gene fragments were purified in accordance with Example III, and digested and sequenced in accordance with Example IV.

10 Of a total of 24 N⁺ (transformed with pBIN19-N⁺) and 18 N⁻ ^{that} (transformed with vector pBIN19) transgenic *N. benthamiana* plants¹ were transferred to soil and grown in the ^{greenhouse, all} greenhouse. All N⁺ lines were confirmed by PCR at leaf stages 4-5 to contain the N gene sequence.

15 The relative level of N protein accumulation was estimated in each independent R₀ transgenic clonal line by DAS-ELISA using antibodies of the TSWV-BL N protein. Of the twenty-four N⁺ lines, two had OD_{405nm} readings of 0.50-1.00, seventeen between 0.02-0.10, and the remaining five less than 0.02. Healthy *N. benthamiana* or transgenic N⁻ plants gave OD_{405nm} readings of 0.00-0.02. All the R₀ plants were self-pollinated and the seeds from the following transgenic lines were
20 germinated on kanamycin (300 µg/ml) selection medium for inoculation tests: (1) N⁻-2 and -6, control transgenic lines containing vector pBIN19 alone; (2) N⁺-28, a transgenic line that produced an undetectable amount of the N protein (OD_{405nm} = 0.005); (3) N⁺-21, a transgenic line producing a low level of the N protein (OD_{405nm} =
25 0.085); and (4) N⁺-34 and -37, two transgenic lines accumulating high levels of the N protein (OD_{405nm} = 0.50-1.00. These six lines were then analyzed by Northern hybridization; the intensity of N gene transcripts correlated well with the levels of ELISA reactions.

30 Transgenic seedlings from the six R₀ lines were selected by germinating seeds on kanamycin selection medium, and these seedlings were inoculated with the five *Tospoviruses*. The inoculated R₁ plants were rated susceptible if virus symptoms were observed on uninoculated leaves. In order to exclude the possibilities of escapes, transgenic control N⁻ plants were always used in each inoculation of
35 transgenic N⁺ plants. In addition, each inoculum extract was always

used to first inoculate N⁺ plants followed by control N⁻ plants. The results from this series of studies are depicted below:

Reactions of R₁ plants expressing the nucleocapsid (N) protein gene of *N. benthamiana* spotted wilt virus (TSWV) to inoculation with *Tospoviruses*

5	No. plants infected/No. plants inoculated ^b						
	R0 Line	ELISA ^a	TSWV ISOLATE		INSV ISOLATE		TSWV-B
			BL	10W	Beg	LI	
	N--2/-6	<0.02	32/32	32/32	32/32	20/20	32/32
	N+-28	0.005	16/16	16/16	15/16		16/16
10	N+-21	0.085	9/40	17/40	39/40	18/20	40/40
	N+-34	0.715	25/28 ^c	28/28	23/28 ^c		28/28
	N+-37	0.510	26/28 ^c	22/22	21/28 ^c	16/20 ^c	22/22

^aELISA data of R₀ lines from which the R₁ plants were derived;

15 ^b30-fold diluted leaf extracts of infected *N. benthamiana* plants were applied to the three leaves of plants at the 3-5 leaf stages. Each extract was always used to inoculate N⁺ plants followed by control N⁻ plants. Data were taken daily for at least two months after inoculation and expressed as number of plants systemically infected/number of plants inoculated;

20 ^cIndicate that nearly all susceptible R₁ plants displayed a significant delay of symptom appearance.

As depicted in the above table, all R₁ plants from control lines N--2 and -6 showed systemic symptoms 5-8 days after inoculation with all the viruses tested. None of the R₁ plants from line N+-28 produced detectable levels of the N protein, and all were susceptible to these viruses except for one plant inoculated with INSV-Beg. ELISA assays of leaf discs from this N+-28 R₁ plant sampled before inoculation clearly showed that the plant identified to possess the INSV-Beg resistant phenotype did accumulate a high level of the N protein (OD_{405nm} = 0.78 as compared to OD_{405nm} <0.02 for all other N+-28 R₁ plants).

35 The low N gene expressing line N+-21 showed the best resistance against the homologous (78%) and closely related TSWV-10W (57%) isolates and very little resistance to the two INSV isolates (3% and 10%); only three N+-21 plants showed the resistant phenotype when inoculated with the INSV isolates. Leaf samples from these INSV-resistant N+-21 R₁ plants gave much higher ELISA reactions (OD_{405nm} 0.5 to 1.00) and thus higher amounts of the N protein than the

100-3410330

susceptible N⁺-21 plants (OD_{405nm} 0.02 to 0.20). The high N gene expressing lines N⁺-34 and -37 showed the highest resistance to INSV isolates (18%-25%) followed by the homologous TSWV-BL isolate (7% and 11%) while none of the plants showed resistance to TSWV-10W; however, the N⁺-34 and -37 R₁ plants that became infected with INSV or TSWV-BL did show various lengths of delays in symptom expression. None of the R₁ plants from these four transgenic N⁺ lines were resistant to TSWV-B; some of the R₁ plants from the N⁺-34 and -37 lines showed a slight delay of symptom appearance.

10 In studies to determine whether the level of N protein production in N⁺ R₁ plants was associated with resistance to different *Tospoviruses*, the inoculated N⁺ R₁ plants in the preceding table were re-organized into four groups based on the intensity of their ELISA reactions of tissues taken before inoculation irrespective of original R₀ plants. The N⁺ R₁ plants that expressed low levels of the N protein (0.02-0.2 OD) showed high resistance (100% and 80%) to TSWV-BL and -10W but were all susceptible to INSV-Beg and -LI, showing no detectable delay in symptom expression relative to control N⁻ plants. In contrast, nearly all N⁺ R₁ plants with high levels of the N protein (0.20-1.00 OD) showed various levels of protection against TSWV-BL, INSV-Beg and -LI, ranging from a short delay of symptom expression to complete resistance with most of these plants showing various lengths of delay in symptom development relative to control N⁻ plants. No protection was observed in the high expressors against TSWV-10W. In addition, none of the N⁺ R₁ plants were resistant to TSWV-B regardless of the level of N gene expression; however, a short delayed symptom appearance was observed in the N⁺ R₁ plants producing high levels of the N protein. All control N⁻ R₁ plants and transgenic N⁺ R₁ plants with undetectable ELISA reactions (0 to 0.02 OD) were susceptible to all the *Tospoviruses* tested.

The inhibition of replication of a distantly related INSV in *N. benthamiana* protoplasts expressing the TSWV-BL nucleocapsid gene was also studied. In these studies, whole INSV-LI virions were used to infect protoplasts that were isolated from three transgenic lines to investigate how the products of the transgene affect replication of the

Incoming virus. Viral replication was determined by measuring the level of the N protein of the infecting INSV in transgenic protoplasts using antibodies specific to the INSV N protein. DAS-ELISA analysis showed that all progenies from a given line were relatively uniform and nearly all R₁ progeny gave an expression level of transgenic N gene similar to their parental transgenic line. These results allowed for the prediction of the expression level of R₁ populations based on that of their parental lines. Protoplasts derived from R₁ plants of the low expressor line N⁺-21 supported the replication of INSV-LI whereas protoplasts from R₁ plants of the higher expressor line N⁺-37 did not until 42 hours after inoculation at which low levels of viral replication were observed. The same protoplasts at various time intervals (e.g. 0, 19, 30 and 42 hours) were also assayed by DAS-ELISA using antibodies specific to the TSWV-BL N protein to monitor the expression level of the transgene. As expected, protoplast from N⁺-21 R₁ plants produced relatively low levels (0.338-0.395 OD_{405nm}) whereas protoplasts from N⁺-37 R₁ plants accumulated high levels (0.822-0.865 OD_{405nm}). The expression level was found to be consistent at all time points.

In this aspect of the present invention it has been shown that transgenic *N. benthamiana* plants that accumulate low amounts of the TSWV-BL N protein are highly resistant to the homologous and closely related (TSWV-10W) isolates, while plants that accumulate high amounts of this protein possess moderate levels of protection against both the homologous and distantly related (INSV-Beg and INSV-LI) viruses. More importantly, these findings indicate that transgenic *N. benthamiana* plants (a systemic host of INSV) are protected against INSV-Beg and INSV-LI isolates.

As discussed above, we have shown that transgenic plants expressing the N gene of TSWV are resistant to homologous isolates, and that such plants expressing the TSWV-BL N gene are resistant to both TSWV and INSV. It has also been shown the best resistance to homologous and closely related isolates was found in transgenic plants accumulating low levels of N protein while transgenic plants with high levels of TSWV-BL N protein were more resistant to serologically distant INSV isolates. This observation led us to suspect the role of

the translated N protein product in the observed protection against homologous and closely related isolates and to speculate that either the N gene itself which was inserted into the plant genome or its transcript was involved in the protection. To test this hypothesis transgenic plants containing the promoterless N gene or expressing the sense or antisense untranslatable N coding sequence were produced. What was discovered was that both sense and antisense untranslatable N gene RNAs provided protection against homologous and closely related isolates, and that these RNA-mediated protections were most effective in plants that synthesized low levels of the respective RNA species and appears to be achieved through the inhibition of viral replication.

More specifically, the coding sequences introduced into transgenic plants is shown in figure 7. As depicted, the construct pBIN19-N contains the promoterless N gene inserted into the plant transformation vector pBIN19 (see Example IV). All other constructs contain a double 35S promoter of CaMV, a 5'-untranslated leader sequence of alfalfa mosaic virus and a 3'-untranslated/polyadenylation sequence of the nopaline synthase gene. pBI525 is a plant expression vector and is used in this study as a control; pBI525-mN contains the mutant (untranslatable) form of the N gene; pBI525-asN contains the antisense form of the untranslatable N gene. One nucleotide deletion at the 5'-terminus of the mutant N gene is indicated by the dash symbol. ATG codons are underlined and inframe termination codons in the mutant gene are shown in bold.

EXAMPLE VIII

Primer-directed mutagenesis and cloning of the TSWV-BL N gene was conducted as follows:

Full-length N gene was obtained by reverse transcription and polymerase chain reaction as described in Phytopathology 82:1223 (1992), the disclosure of which is incorporated *in toto* herein. The untranslatable N coding sequence was similarly generated by RT-PCR using oligomer primers A (AGCATTGGATCCATGGTTAACACACTAAGCAAGCAC), which is identical to the S RNA in the 3'-noncoding region of the TSWV-BL N gene, and B (AGCTAATCTAGAACCATGGATGACTCACTAAGGAAAGCATTGTTGC),

complementary to the S RNA in the 5'-terminus of the N gene. The latter oligomer primer contains a frameshift mutation immediately after the translation initiation codon and several termination codons to block possible translation readthroughs. The intact and mutant N gene

5 fragments were purified on a 1.2% agarose gel as described in Example II. The gel-isolated intact and mutant N gene fragments were digested with the appropriate restriction enzyme(s) and directly cloned into

BamHI/XbaI-digested plant transformation vector pBIN19 and NcoI-

10 digested plant expression vector pBI525, respectively as described in Example IV. The resulting plasmids were identified and designated as pBIN19-N containing the intact, promoterless N gene, and pBI525-mN and pBI525-asN containing the mutant coding sequence in the sense and antisense orientations, respectively, relative to cauliflower mosaic virus 35S promoter. The translatability of the mutant N coding

15 sequence in the expression cassette was checked by transient expression assay in *Nicotiana tabacum* protoplasts; and the expression cassettes containing the sense or antisense mutant N coding sequence were then excised from plasmid pBI525 by a partial digestion with

20 HindIII/EcoRI (since the N coding sequence contains internal HindIII and EcoRI sites), and ligated into the plant transformation vector pBIN19 that had been cut with the same enzymes. The resulting vectors as well as pBIN19-N were transferred to *A. tumefaciens* strain LBA4404 using the procedure described in Example IV. Leaf discs of *N. tabacum* var

Havana cv 423 were inoculated with the *A. tumefaciens* strain LBA4404

25 containing various constructs and the resulting transgenic plants were self-pollinated and seeds were selectively germinated on kanamycin medium.

PCR was performed on each R₀ transgenic line as described above. The oligomer primers A and B were used to determine the
30 presence of the N coding sequence of TSWV-BL. The oligomer primer 35S-promoter (CCCACTATCCTTCGCAAGACCCY^(SEQ ID NO. 24)) was combined with either the oligomer primer A or B to confirm the orientation (relative to the CaMV 35S promoter) of the mutant N coding sequence inserted into the plant genome. DAS-ELISA used to detect the N protein in transgenic
35 plants was performed using polyclonal antibodies against the TSWV-BL

N protein. For an estimation of RNA transcript level in transgenic plants by Northern blot, total plant RNAs were isolated according to Napoli [see The Plant Cell 2:279 (1990)], and were separated on a formaldehyde-containing agarose gel (10 µg/lane). The agarose gels were then stained with ethidium bromide to ensure uniformity of total plant RNAs in each lane. Hybridization conditions were as described in the GeneScreen Plus protocol by the manufacturer.

Resulting signal blots were compared and normalized based on the N gene transcript band of the control lane (the mN R₁ plant producing a high level of the N gene transcript) included in each blot. The transgenic plants that gave density readings (Hewlett ScanJetTM and Image Analysis Program) between 100 and 150 were rated as high expressors, while the plants with densities between 15 and 50 were rated as low expressors.

Inoculation of transgenic plants with *Tospovirus* was carried out as described above with inoculation being done at the 3-4 leaf stage except were indicated.

Tobacco protoplasts were prepared from surface-sterilized leaves derived from R₁ plants [see Z. Pflanzanphysiol. 78:453 (1992) with modifications]. The isolated protoplasts (6×10^6 protoplasts) were transformed with 0.68 OD_{260nm} of the purified TSWV-BL virion preparation using the PEG method [see Plant Mol. Biol. 8:363 (1987)]. The transformed protoplasts were then cultured at the final density of 1×10^6 protoplasts /ml in the culture medium at 26°C in the dark.

After various intervals of incubation, the cultured protoplasts were washed twice with W5 solution and lysed by osmotic shock in the enzyme conjugate buffer. Viral multiplication (replication) was estimated by measuring the N protein of the virus using DAS-ELISA.

As described, one aspect of the present invention demonstrated that transgenic tobacco producing none or barely detectable amounts of the N protein were resistant to homologous and closely related isolates. This result suggested that the observed resistance may have been due to trans interactions of the incoming viral N gene RNA with either the N gene transcript produced in the transgenic plants or the N coding sequence itself. To test whether the presence of the nuclear N gene

plays a role, transgenic P^oN R₀ lines and R₁ plants from two P^oN lines were challenged with four *Tospoviruses* (TSWV-BL, TSWV-10W, INSV-Beg and TSWV-B). Only asymptomatic plants were rated resistant while plants showing any symptoms were rated susceptible. All inoculated R₀ and R₁ plants were susceptible to the viruses.

To further test the possibility that the transcript of the N transgene is involved in the protection, a number of R₀ transgenic plants that produced either the sense or the antisense N gene transcript but not the N protein were inoculated with the homologous isolate.

Results appear in the following table:

Form of transgene ^a	Level of N gene RNA ^b	No. of R ₀ lines tested	No. of lines inoculated ^c	No. of lines resistant
mN	H	8	4	0
	L	17	16	16
asN	nd	4	1	0
	H	6	3	0
	L	9	5	5
	nd	1	0	0
P ^o N	nd	12	6	0

^amN and asN represent plants expressing the sense and antisense untranslatable N genes; respectively, P^oN represents plants containing the promoterless N gene;

^bthe level of the N gene RNA was estimated in each line by Northern blots, nd indicates that the N gene transcript was not detected;

^c30-fold diluted leaf extracts of the *N. benthamiana* plants infected with TSWV-BL were applied to three leaves of each plant at the 6-7 leaf stage. Each extract was first applied to all test plants followed by control healthy plants. Data were taken daily for 45 days after inoculation and only the asymptomatic plants were rated resistant.

Unlike the controls, which developed typical systemic symptoms 7 to 9 days after inoculation, 16 out of 21 mN plants and 5 out of 8 asN plants were asymptomatic throughout their life cycles. Northern blot analysis of leaf tissues sampled before inoculation showed that all the resistant R₀ lines produced low levels of the sense or antisense N gene RNA, whereas the susceptible R₀ lines produced either none or high levels of the RNA species. Since this data suggested that the resistance of transgenic plants to TSWV-BL was related to their relative levels of N gene transcript, transgenic progenies from four mN

and three asN R₀ lines with either high or low N gene transcript levels were selected by germination on kanamycin-containing media. These transgenic plants were tested for resistance to the four *Tospoviruses* at the 3 to 4 leaf stage, except that some R₁ plants from two asN lines
5. were inoculated at the 6 to 7 leaf stage. The results are summarized in the following table:

<u>R₀ Line</u>	<u>N gene RNAa</u>	<u>ISWV-BL</u>	<u>ISWV-10W</u>	<u>INSV-Beg</u>	<u>ISWV-B</u>
<u>Promoterless N gene</u>					
P ^o N-1	nd	10/10	10/10	10/10	10/10
P ^o N-2	nd	15/15	10/10	10/10	10/10
N ^o -3	nd	8/8	6/6	6/6	6/6
<u>Untranslatable N gene</u>					
mN-2	H	20/20	20/20	20/20	20/20
mN-7	H	20/20	20/20	20/20	20/20
mN-13	L	2/20	4/20	20/20	20/20
mN-18	L	4/20	1/20	20/20	20/20
N ^o -3	nd	24/24	32/32	24/24	24/24
<u>Antisense N gene</u>					
asN-1	L	<u>20/20</u> ^b	<u>20/20</u>	20/20	20/20
asN-4	H	20/20 (16/16) ^c	20/20 (16/16)	20/20	20/20
asN-9	L	<u>19/20</u> (3/41)	<u>20/20</u> (5/21)	20/20	20/20
N ^o -3	nd	16/16 (32/32)	16/16 (20/20)	16/16	16/16

^aNorthern analysis of R₀ lines from which the R₁ plants were derived (see preceding table);

^bthe underlined fractions indicate that most of susceptible R₁ plants displayed a significant delay of symptom appearance;

^cthe fraction in parenthesis represents the inoculation data obtained from plants inoculated at the 6-7 leaf stage; the remaining data in this table were generated from plants inoculated at the 3-4 leaf stage; inoculated plants were observed daily for 45 days after inoculation.

All R₁ plants from high expressor lines mN-2 and mN-7 were susceptible to infections by all *Tospoviruses* tested, and these plants did not show a delay of symptom appearance as compared to controls. In contrast, high proportions of the R₁ plants from low expressor lines mN-13 and -18 were resistant to homologous (TSWV-BL) and closely related (TSWV-10W) isolates, but not resistant to infections by distantly related *Tospoviruses* (INSV-Beg and TSWV-B). The resistance of asN R₁ plants from low expressor R₀ lines was markedly influenced by the TSWV isolate used for inoculation. All but one of the small R₁ plants (3-4 leaf stage) from low expressor lines asN-1 and -9 became infected, although there was a delay of symptom appearance, when inoculated with the homologous TSWV-BL or closely related TSWV-10W isolates. In contrast, most of the large R₁ plants (6-7 leaf stage) from line asN-9 were resistant to both isolates. In comparison, control R₁ plants and R₁ plants from the high expressor line such as asN-4 displayed no resistance to either of the isolates regardless of the size of test plants. Antisense RNA-mediated protection was not effective against infection by the distantly related INSV-Beg and TSWV-B isolates.

Analyses of data presented in the above two tables suggest that sense and antisense RNA-mediated protections are observed only in low expressors of the N gene. The R₁ asN plants that produced high levels of the antisense N gene transcript were as susceptible as control plants. In contrast, the asN low expressors displayed a delay in symptom appearance when inoculated at the 3-4 leaf stage and showed increased levels of resistance when inoculated at the 6-7 leaf stage.

Inhibition of viral replication in tobacco protoplasts expressing the sense or antisense form of untranslatable N coding sequence was also noted. In this instance, whole virion preparations of TSWV-BL were used to transfect protoplasts isolated from transgenic lines to investigate the effect of sense or antisense N gene transcript on replication of the incoming virus. Viral replication was determined by measuring the level of the N protein of the incoming virus in transfected protoplasts, and it was found that protoplasts derived from plants (mN-7 and asN-4) that produced high levels of the respective

RNA transcripts supported the replication of the virus, whereas protoplasts from mN low expressor (mN-18) did not. Protoplasts from an asN low expressor (asN-9) supported much lower levels of viral replication.

Accordingly, in this aspect of the present invention we have shown that transgenic plants expressing sense or antisense form of untranslatable N gene coding sequence are resistant to homologous (TSWV-BL) and closely related (TSWV-10W), but not to distantly related (INSV-Beg and TSWV-B) *Tospoviruses*. The following table provides a comparison of resistance to *Tospoviruses* between transgenic tobacco expressing various forms of the TSWV-BL N gene:

<u>Tospovirus</u>	Homology to TSWV-BL N Gene ^b	<u>Form of the Transgene^a</u>			
		N	mN	asN	P ^o N
TSWV-BL	100%	R	R	RC	S
TSWV-10W	99%	R	R	RC	S
INSV-Beg	60%	RC	S	S	S
TSWV-B	78%	S	S	S	S

Reactions of transgenic tobacco and *N. benthamiana* plants expressing the intact N gene (N) of TSWV-BL to inoculation with the four *Tospoviruses* are included for comparisons with inoculation results of transgenic plants containing untranslatable (mN), antisense (asN), and promoterless (P^oN) N coding sequences, R = resistant, S = susceptible;

^bthe nucleotide sequences are as reported in Phytopathology 82:1223 (1992) and Phytopathology 83:728 (1993)

^clevel of resistance may depend upon the concentration of inoculum.

These results confirm and extend the earlier aspects of the present invention for RNA-mediated protection with TSWV.

Furthermore, the protection is observed in plants producing low rather than high levels of the N gene transcript, and although earlier studies reported herein indicate that tobacco plants which produced high levels of the TSWV-BL N protein displayed resistance to INSV-Beg, this additional data indicates that since resistance to INSV-Beg was not observed in transgenic plants expressing the sense or antisense form of the untranslatable of the N gene thus clearly indicating that protection against INSV-Beg is due to the presence of the N protein and not the N gene transcript. Thus, it appears that two different mechanisms are

involved in protection transgenic plants against TSWV and INSV *Tospoviruses* according to the present invention. One mechanism involves the N gene transcript (RNA-mediated), and another involves the N protein (protein-mediated). In addition, the results of the protoplast experiments indicate that N gene RNA-mediated protection is achieved through a process that inhibits viral replication, and the data contained in the above tables suggest that protection against the distantly related INSV-Beg isolate is conferred by the N protein of TSWV-BI, and not by the gene transcript.

Finally, further studies were conducted to provide still another aspect of the present invention - that a portion of the ^{All} *Tospovirus* nucleoprotein gene provide protection of transgenic plants against infection by the *Tospovirus*. It has already been demonstrated above, that the N gene RNA protects against homologous and closely ^{related} TSWV isolates while the N protein protects against the homologous isolate and distantly related INSV isolates; that N gene ^{RNA} ~~RNA~~-mediated protection is effective in plants expressing low levels of the N gene whereas N protein-mediated protection requires high levels of N protein accumulation; and that the N gene RNA-mediated protection is achieved through inhibition of viral replication. Based upon this prior data, we next set out to determine whether a portion of the N gene might work against infection by the virus. We found, as discussed below, that transgenic plants expressing about one-half of the N gene sequence is resistant to the virus.

The following describes the cloning of one-half N gene fragments of TSWV-BL in order to demonstrate this final aspect of the present invention. The first and second halves of both the translatable and untranslatable N gene were ^{generated} ~~generated~~ by reverse transcription and then PCR as described above. As depicted in figure 8, the nucleotide deletion or insertions at the 5'-terminals of the untranslatable half N gene fragments are indicated by the dash symbol; ATG codons are underlined and all possible termination codons immediately after the initiation codon of the untranslatable half N gene fragments are shown in bold.

The first half of the N gene was produced by RT-PCR using (SEQ ID NO. 25) oligoprimers I (5'-TACAGTGGATCCATGGTTAAGGTAATCCATAGGCTTGAC), which is complementary to the central region of the TSWV-BL N gene, and II (5'-AGCTAAGGATGGTTAAGCTCACTAAGGAAAGCATTGTTGC) for the (SEQ ID NO. 26) translatable or III (5'-AGGTAATOTAGAAGCATGGATGACTCACTAAGGAAAGCATTGTTGC) for (SEQ ID NO. 27) the untranslatable first half N gene fragment, the latter two oligomer primers are identical to the 5'-terminus of the N gene. Similarly, the second half of the N gene was produced by RT-PCR using oligomer (SEQ ID NO. 28) primers IV (5'-AGCATTGGATCCATGGTTAACACACTAAGCAAGCAC) which is complementary to the 3'-noncoding region of the TSWV-BL N gene, and V (5'-TACAGTTCTAGAACCATGGATGATGCAAAGTCTGTGAGG) for the (SEQ ID NO. 29) translatable or VI

(5'-AGATTCTCTAGACCATGGTGACTTGATGAGCAAAGTCTGTGAGGCTTGC)

(SEQ ID NO. 30) ^{Untranslatable Second} for the ~~untranslatable~~ second half N gene fragment, the latter two oligomer primers are identical to the central region of the N gene. The oligomer primer III contains a frameshift mutation immediately after the translation codon and several termination codons to block possible translation readthroughs while the oligomer primer VI contains several inframe termination codons immediately after the translation initiation codon.

The half gene fragments were purified on a 1.2% agarose gel as described above, and the gel-isolated gene fragments were digested with the restriction enzyme *NcoI* and directly cloned into *NcoI*-digested plant expression vector pBI525. The resulting plasmids were identified and designated as (1) pBI525-1N containing the first half translatable N gene, (2) pBI525-1N' containing the first half untranslatable N gene, (3) pBI525-1N⁻ containing the first half translatable N gene in the antisense orientation, (4) pBI525-2N containing the second half translatable N gene, (5) pBI525-2N' containing the second half untranslatable N gene, and (6) pBI525-2N⁻ containing the second half translatable N gene in the antisense orientation. The ^{expression} ~~expression~~ cassettes were then excised from plasmid pBI525 by digestion with *HindIII*/*EcoRI* and ligated as described above into the plant transformation vector pBIN19 that had been cut with the

same enzymes. The resulting vectors as well as plasmid pBIN19 were transferred to *A. tumefaciens* strain LBA4404, using the procedure described by Holsters *supra*. Leaf discs of *N. benthamiana* were inoculated with *A. tumefaciens* strain LBA4404 containing the various constructs. Transgenic plants were self-pollinated and seeds were selectively germinated on kanamycin as described above.

Analysis of transgenic plants by PCR and Northern hybridization PCR was performed on each R₀ transgenic line as described previously. The oligomer primers I to VI were used to determine the presence of the N coding sequence of TSWV-BL. The oligomer primer 35S-Promoter (see Example VIII) was combined with one of the above oligomer primers to confirm the orientation (relative to the CaMV 35S promoter) of the half gene sequences inserted into the plant genome. Northern analysis was conducted as described in Example VIII.

Lettuce isolate of TSWV (TSWV-BL) was used to challenge transgenic plants. Inoculation was done using test plants at the 3-4 leaf stage as described above. To avoid the possibility of escapes, control plants were used in each experiment and each inoculum extract was used to first inoculate the transgenic plants followed by control plants.

The various constructs used in this aspect of the present invention are illustrated in figure 8. Translatable and untranslatable half N gene ^{fragments} were synthesized by RT-PCR and then cloned directly into the plant expression vector pBI525. The oligomer primers iii and vi, used for generation of untranslatable half N gene fragments by RT-PCR, contains a mutation immediately after the translation initiation codon and the resulting reading frame contains several termination codons to block possible translation readthroughs. Thus, both first and second half untranslatable N gene fragments should be incapable of ^{producing} the truncated N protein fragments when introduced into plants. Both translatable and untranslatable half N gene ^{fragments} were then placed downstream of the CaMV 35S promoter of the vector pBI525 in the sense orientation or in the antisense orientation. The ^{expression} of the half N coding sequences of TSWV-BL was thus controlled by a double CaMV 35S promoter fused to the 5'-

untranslated leader sequence of alfalfa mosaic virus (ALMV) of the expression vector pBI525. Expression vectors that utilize the stacked double CaMV 35S promoter elements are known to yield higher levels of mRNA ^{transcription} than similar vectors with a single 35S promoter element. Expression cassettes were transferred from the vector pBI525 to the ^{plant} transformation vector pBIN19. The ^{resulting} plasmids as well as the control plasmid pBIN19 were then transferred into *A. tumefaciens* strain LBA4404. Transgenic plants were obtained with nomenclature of the transgenic lines shown in figure 8.

All the kanamycin-resistant transgenic lines were confirmed by PCR to contain the proper N coding sequences in the expected orientations. Each transgenic R₀ line which was grown for seeds was then assayed using Northern blot. Six out of six 1N, four out of six 1N', six out of six 1N⁻, six out of six 2N, seven out of eight 2N', and six out of seven 2N⁻ transgenic R₀ lines were found to produce half N gene RNAs.

A set of transgenic R₀ plants was challenged with the homologous isolate TSWV-BL. Only asymptomatic plants were rated resistant while the plants showing any symptom (local lesions or systemic infections) were rated susceptible. All the inoculated R₀ control plants were susceptible to the virus; in contrast, two out of nine 1N, two out of six 1N⁻, four out of ten 2N', and one out of eight 2N⁻ R₀ lines were found to be completely resistant to the virus infection. Although none of the 1N and 2N R₀ lines showed high levels of resistance, some of those plants displayed significant delays of symptom appearance.

Another set of transgenic R₀ lines was brought to maturity for seed production. Seedlings were germinated on kanamycin-containing medium and inoculated with TSWV-BL. As shown in the following table, control seedlings and seedlings from some of the transgenic lines were susceptible to the isolate whereas seedlings from lines 1N-151, 1N⁻-123, and 2N⁻-134 showed ^{various} levels of protection, ranging from delays in symptom expression to complete resistance.

	R0 line	<u>No. plants infected/No. plants inoculated</u>		
		6DPI	15DPI	30DPI
	Control	50/50		
	1N-149	17/17		
5	1N-151	2/20	13/20	17/20
	1N-123	16/20	17/20	17/20
	1N-124	20/20		
	1N-126	19/19		
	1N-130	12/15	15/15	
10	1N-132	18/19	19/19	
	2N-155	20/20		
	2N-134	0/20	10/20	10/20
	2N-135	19/19		
	2N-142	20/20		
15	2N-143	20/20		

In the above table, 30-fold diluted extracts of infected *N. benthamiana* were used to inoculate transgenic plants at the 3-4 leaf stage followed by control transgenic plants. DPI = days post inoculation.

In summary, this aspect of the present invention shows that transgenic plants expressing the first or the second half of either translatable or untranslatable N gene fragment are highly resistant to the homologous TSWV-BL isolate. This result demonstrates that a portion of the N gene is sufficient for resistance to the virus.

A listing of all nucleotide and amino acid sequences described in the foregoing description of the present invention is as follows:

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Dennis Gonsalves and Sheng-Zhi Pang

(ii) TITLE OF INVENTION: Tomato Spotted Wilt Virus

(iii) NUMBER OF SEQUENCES: 30

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGCAGGCAAA ACTGCGAGAA CTTCG 25

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCAAGTTCG CGAGTTTTCG CTGCT 25

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCTAACCAT GGTAAAGCTC ACTAAGGAAA GC 32

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGCATTCAT GGTAAACACA CTAAGCAAGC AC 32

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2265 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CAAGTTGAAA GCAACAACAG AACTGTAAAT TCTCTTGCAG TGAAATCTCT 50
GCTCATGTCA GCAGAAAACA ACATCATGCC TAACTCTCAA GCTTCCACTG 100
ATTCTCATT CAAGCTGAGC CTCTGGCTAA GGGTTCCAAA GGTTTTGAAG 150

SUB A6
cont

66200T 60492160

	CAGGTTTCCA TTCAGAAATT GTTCAAGGTT GCAGGAGATG AAACAAACAA	200
	AAATTTTTAT TTATCTATTG CCTGCATTCC AAACCATAAC AGTGTTGAGA	250
	CAGGTTTAAA CATTACTGTT ATTTGCAAGC ATCAGCTOCC AATTGCGAAA	300
	TGCAAAAGCTC GTTTTGAATT ATCAATGATG TTTTCTGATT TAAAGGAGCC	350
5	TTACAACATT GTTCATGACC GTTCATAACC CAAAGGATCG GTTCCAATGC	400
	TCTGGGTGCA AACTCAGACA TCTTTGCACA AGTTCTTTGC AACTAACTTG	450
	CAAGAAGATG TAATCATCTA CACTTTGAAC AACCTTGAGC TAACTCCTGG	500
	AAAGTTAGAT TTAGGTGAAA GAACCTTGAA TTACAGTGAA GATGCTTACA	550
	AAAGGAAATA TTTCTTTTCA AAAACACTTG AATGCTTCC ATCTAACACA	600
10	CAAACATGT CTTACTTGA CAGCATCCAA ATCCCTTCAT GGAAGATAGA	650
	CTTTGCCAGA GGAGAAATTA AAATTTCTCC ACAATCTATT TCAGTTGCAA	700
	AATCTTTGTT AAAGCTTGAT TTAAGOGGGA TCAAAAAGAA AGAATCTAAG	750
	GTTAAGGAAG CGTATGCTTC AGGATCAAAA TAATCTTGCT TTGTCCAGCT	800
	TTTTCTAATT ATGTTATGTT TATTTTCTTT CTTACTTAT AATTATTCT	850
15	CTGTTTGICA TCCTTTTCAA ATTCTCTCTG TCTAGTAGAA ACCATAAAAA	900
	CAAAAAATAA AAATGAAAAT AAAATTAATA TAAATAAAAA TCAAAAAATG	1000
	AAATAAAAAC AACAAAAAAT TAAAAAAGA AAAACCAAAA AGACCCGAAA	1050
	GGGACCAATT TGGCCAAATT TGGGTTTTGT TTTTGTTTTT TGTTTTTTGT	1100
	TTTTTATTTT TTATTTTATT TTTATTTTAT TTTATTTTAA TTTTATTTT	1150
20	ATTTTATTTA TTTTTTGTTT TCGTTTGTTT TGTATTTTAA TTATTTATTA	1200
	AGCACAACAC ACAGAAAGCA AACTTTAATT AAACACACTT ATTTAAAATT	1250
	TAACACACTA AGCAAGCACA AGCAATAAAG ATAAAGAAAG CTTTATATAT	1300
	TTATAGGCTT TTTTATAATT TAACTTACAG CTGCTTTCAA GCAAGTTCTG	1350
	CGAGTTTTGC CTGCTTTTAA ACCCOGAACA TTTCATAGAA CTGTGTAAGA	1400
25	GTTCACCTGT AATGTTCCAT AGCAACACTC CCTTTAGCAT TAGGATTGCT	1450
	GGAGCTAAGT ATAGCAGCAT ACTCTTTCC CTTCTTCACC TGATCTTCAT	1500
	TCATTTCAA TGCTTTGCTT TFCAGCACAG TGCAAACTTT TCTAAGGCT	1550
	TCCTTGGTGT CATACTTCTT TGGGTCGATC COGAGGTCT TGTATTTTGC	1600
	ATCCTGATAT ATAGCCAAGA CAACACTGAT CATCTCAAAG CTATCAACTG	1650
30	AAGCAATAAG AGGTAAGCTA CCTCCAGCA TTATGGCAAG TCTCACAGAC	1700
	TTTGATCAT CGAGAGGTA TCCATAGGCT TGAATCAAAG GATGGGAGC	1750
	AATCTTAGAT TTGATAGTAT TGAGATTCTC AGAATTCCA GTTCTTTCAA	1800
	CAAGCCTGAC CCTGATCAAG CTATCAAGCC TTCTGAAGGT CATGTCAGTG	1850
	CCTCCAATCC TGTCTGAAGT TTTCTTTATG GTAATTTTAC CAAAAGTAAA	1900
35	ATCGCTTTGC TTAATAACCT TCATTATGCT CTGACGATT TTIAGGAATG	1950

TCAGACATGA AATAACGCTC ATCTTCTTGA TCTGGTCGAT GTTTTCCAGA 2000
 CAAAAAGTCT TGAAGTTGAA TGCTACCAGA TTCTGATCTT CCTCAAACCTC 2050
 AAGGCTCTTG CCTTGIGTCA ACAAAGCAAC AATGCTTTCC TTAGTGAGCT 2100
 TAACCTTAGA CATGATGATC GTAAAAGTTG TTATAGCTTT GACCGTATGT 2150
 5 AACTCAAGGT GCGAAAGTGC AACTCTGTAT CCGCAGTGC TTTCTTAGGT 2200
 TCTTAATGIG ATGATTGTGA AGACTGAGTG TTAACGTATG AACACAAAAT 2250
 TGACAGATT GCTCT 2265

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1709 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AAATTCTCTT GCAGTGAAT CTCIGCTCAT GTTAGCAGAA AACAACATCA 50
 TGCCTAACTC TCAAGCTTTT GTCAAGCTT CTAATGATTC TAATTTCAAG 100
 CTGAGCCTCT GGCTAAGGGT TCCAAAGGTT TTGAAGCAGA TTTCATTCA 150
 GAAATTGTTT AAGGTTGCAG GAGATGAAAC AAATAAAACA TTTTATTTAT 200
 20 CTATTGCCTG CATTCCAAAC CATAACAGTG TTGAGACAGC TTAAACATT 250
 ACTGTTATTT GCAAGCATCA GCTOCCAATT CGTAAATGTA AACTCCTTT 300
 TGAATTATCA ATGATGTTTT CTGATTTAAA GGAGCCTTAC AACATTATTC 350
 ATGATCCTTC ATATCCCAA AGGATTGTTT ATGCTCTGCT TGAAACTCAC 400
 ACATCTTTTG CACAAGTCTT TTGCAACAAC TTGCAAGAAG ATGTGATCAT 450
 25 CTACACCTTG AACAAOCATG AGCTAACTCC TGGAAAGTGA GATTTAGGIG 500
 AAATAACTTT GAATTACAAT GAAGACGCTT ACAAAGGAA ATATTTCCCT 550
 TCAAAAACAC TTGAATGTCT TCCATCTAAC ATACAAACTA TGTCTTATTT 600
 AGACAGCATC CAAATOCCTT OCTGGAAGAT AGACTTTGCC AGGGGAGAAA 650
 TTAAATTTT TCCACAATCT ATTTTCAGTTG CAAATCTTT GTTAAATCTT 700
 30 GATTTAAGCG GGATTAAAAA GAAAGAATCT AAGATTAAGG AAGCATATGC 750
 TTCAGGATCA AAATGATCTT GCTGTGTCCA GCCTTTTCTA ATTATGTTAT 800
 GTTATTTTTC TTTCTTACT TATAATTATT TTTCTGTTTG TCATTCTTT 850
 CAAATTCCTC CTGTCTAGTA GAAACCATAA AAACAAAAAT AAAAATAAAA 900
 TAAATCAAAA ATAAATAAAA AATCAAAAAA TGAAATAAAA GCAACAAAAA 950
 35 AATTAAAAAA CAAAAACCA AAAAAGATCC CGAAAGGACA ATTTTGGCCA 1000
 AATTGGGGT TTGTTTTGT TTTTGTGTTT TTGTTTTT GTTTTATTT 1050

SUB A6
cont 10

	Leu	Glu	Thr	His	Thr	Ser	Leu	His	Lys	Phe	Phe	Ala	Thr	Asn	Leu
					140					145					150
	Gln	Glu	Asp	Val	Ile	Ile	Tyr	Thr	Leu	Asn	Asn	Leu	Glu	Leu	Thr
					155					160					165
5	Pro	Gly	Lys	Leu	Asp	Leu	Gly	Glu	Arg	Thr	Leu	Asn	Tyr	Ser	Glu
					170					175					180
	Asp	Ala	Tyr	Lys	Arg	Asp	Tyr	Phe	Leu	Ser	Lys	Thr	Leu	Glu	Cys
					185					190					195
	Leu	Pro	Ser	Asn	Thr	Gln	Thr	Met	Ser	Tyr	Leu	Asp	Ser	Ile	Gln
					200					205					210
	Ile	Pro	Ser	Tyr	Lys	Ile	Asp	Phe	Ala	Arg	Gly	Glu	Ile	Lys	Ile
					215					220					225
	Ser	Pro	Gln	Ser	Ile	Ser	Val	Ala	Lys	Ser	Leu	Leu	Lys	Leu	Asp
					230					235					240
15	Leu	Ser	Gly	Ile	Lys	Lys	Lys	Glu	Ser	Lys	Val	Lys	Glu	Ala	Tyr
					245					250					255
	Ala	Ser	Gly	Ser	Lys										
					260										

(2) INFORMATION FOR SEQ ID NO:8:

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 858 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

	TTAACACACT	AAGCAAGCAC	AAACAATAAA	GATAAAGAAA	GCTTTATATA	50
	TTTATAGGCT	TTTTTATAAT	TTAACTTACA	GCTGCTTTTA	AGCAAGTTCT	100
	GTGAGTTTIG	CCTGTTTTTT	AACCCCAAAC	ATTTCATAGA	ACTTGTTAAG	150
30	GGTTTCACIG	TAATGTTCCA	TAGCAATACT	TCCITTAGCA	TTAGGATTGC	200
	TGGAGCTAAG	TATAGCAGCA	TACTCTTTCC	CCTTCCTCAC	CTGATCTTCA	250
	TTCATTTCAA	ATGCTTTTCT	TTTCAGCACA	GIGCAAACIT	TTCCTAAGGC	300
	TTCCCTGGIG	TCATACTTCT	TTGGGTGCGAT	CCCGAGATCC	TGTATTTTIG	350
	CATCCIGATA	TATAGCCAAG	ACAACACIGA	TCATCTCAAA	GCTATCAACT	400
35	GAAGCAATAA	GAGGTAAGCT	ACCTCCAGC	ATTATGGCAA	GCCACACAGA	450
	CTTTGCATCA	TCAAGAGGTA	ATCCATAGGC	TTGACTCAAA	GGGTGGGAAG	500
	CAATCTTAGA	TTTGATAGTA	TTGAGATTCT	CAGAATTCCC	AGTTTCTCTA	550
	ACAAGCCTGA	CCCTGATCAA	GCTATCAAGC	CTTCTGAAGG	TCATGTGAGT	600
	GGCTCCAATC	CTGTCTGAAG	TTTCTTTTAT	GGTAATTTTA	CCAAAAGTAA	650
40	AATCGCTTTG	CTTAATAAOC	TTCATTATGC	TCTGACGATT	CTTCAGGAAT	700

GTCAGACATG AAATAATGCT CATCTTTTIG ATCTGGTCAA GGTTTTCCAG 750
 ACAAAGAGTC TTGAAGTTGA ATGCTACAG ATTCTGATCT TCCTCAAAC 800
 CAAGGCTTT GCCTTGIGTC AACAAAGCAA CAATGCTTTC CTAGTGAGC 850
 TTAACAT 858

5 (2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2028 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AAATTCCTCTT GCAGTCAAT CTCTGCTCAT GTTAGCAGAA AACAAATCA 50
 TGCCTAAGTC TCAAGCTTTT GTCAAAGCTT CTACTGATTC TAATTTCAAG 100
 15 CTGAGCCTCT GGCTAAGGT TOCAAAGGT TIGAAGCAGA TTTCCATCA 150
 GAAATTGTTT AAGGTTGCAG GAGATGAAAC AAATAAAACA TTTTATTTAT 200
 CTATTGCTTG CATTOCAAAC CATAACAGTG TTGAGACAGC TTTAAACATT 250
 ACTGTTATTT GCAAGCATCA GCTCCCAATT CGTAAATGTA AAACCTCTTT 300
 TGAATTATCA ATGATGTTTT CTGATTTAAA GGAGCCTTAC AACATTATTC 350
 20 ATGATCCCTC ATATCCCAAA AGGATGTTTC ATGCTCTGCT TGAAACTCAC 400
 ACATCTTTTG CACAAGTTCT TTGCAACAC TTGCAAGAAG ATGIGATCAT 450
 CTACACCTTG AACAAACATG AGCTAACTTC TGGAAAGTTA GATTTAGGTG 500
 AAATAACTTT GAATTACAAT GAAGACGCT ACAAAGGAA ATATTTCTCT 550
 TCAAAAACAC TTGAATGTTT TOCATCTAAC ATACAAACTA TGCTTATTT 600
 25 AGACAGCATC CAAATCCCTT CCTGGAAGAT AGACTTTGCC AGGGGAGAAA 650
 TTAATAATTT TCCACAATCT ATTTGAGTTG CAAATCTTT GTTAAATCTT 700
 GATTTAAGCG GGATTAAAAA GAAAGAATCT AAGATTAAAG AAGCATATGC 750
 TTCAGGATCA AAATGATCTT GCTGIGTCA GCTTTTTCTA ATTATGTTAT 800
 GTTATTTTTC TTCTTTACT TATAATTATT TTTCTGTTG TCATTTCTTT 850
 30 CAAATCTCTC CTGCTAGTA GAAACATAA AAACAAAAAT AAAAATAAAA 900
 TAAATCAAA ATAAATAAA AATCAAAAA TGAAATAAAA GCAACAAAA 950
 AATTAAAAA CAAAAACCA AAAAGATCC CGAAAGGACA ATTTTGGCCA 1000
 AATTGGGGT TIGTTTTGT TTTTGTGTTT TTGTTTTT GTTTTATTT 1050
 TTATTTTAT TTTTATTTT ATTTTATTT ATTTATGTT TTTGTTGTT 1100
 35 TIGTTATTT GTTATTTAT AAGCACAACA CACAGAAAGC AAATTTAAT 1150
 TAAACACACT TATTTAAAT TTAACACACT AAGCAAGCAC AAACAATAA 1200

SUB A6
CONT

662207-03492160

GATAAAGAAA GCTTTATATA TTTATAGGCT TTTTATAAT TTAACCTACA 1250
 GGTGGTTTTA AGCAAGTTCT GIGAGTTTTG CCTGTTTTTT AACCCCAAAC 1300
 ATTTCATAGA ACTTGTAAAG GGTTCACCTG TAATGTTCCA TAGCAATACT 1350
 TCGTTTAGCA TTAGGATTGG TGGAGCTAAG TATAGCAGCA TACTCTTTCC 1400
 5 CCTTCTTACG CIGATCTTCA TTCAATTTCA ATGCTTTTCT TTTCAGCACA 1450
 GGGCAAACCT TTCCCTAAGGC TTCCCTGGTG TCATACTTCT TTGGGTGGAT 1500
 CCGGAGATCC TTGTATTTTG CATCCGTGATA TATAGCCAAG ACAACACTGA 1550
 TCATCTCAAA GGTATCAACT GAAGCAATAA GAGGTAAGCT AACTCCAGC 1600
 ATTATGGCAA GCTCAGAGA CTTCGCATCA TCAAGAGGTA ATCCATAGGC 1650
 10 TTGACTCAA GGGTGGGAAG CAATCTTAGA TTGATAGTA TTGAGATTCT 1700
 CAGAATTCCC AGTTTCCCTCA ACAAGCCTGA CCTGATCAA GCTATCAAGC 1750
 CTTCTGAAGG TCATGTCAGT GGCTCCAATC CTGTCGAAG TTTTCTTTAT 1800
 GGTAAATTTA CCAAAGTAA AATCGCTTTG CTTAATAACC TTCATTATGC 1850
 TCTGACGATT CTTCAAGGAAT GTCAGACATG AAATAATGCT CATCTTTTGG 1900
 15 ATCTGGICAA GGTTTTCCAG ACAAAAAGTC TTGAAGTTGA ATGCTACCAG 1950
 ATTCTGATCT TCCTCAAACT CAAGGCTTTT GCTTGTGTC AACAAAGCAA 2000
 CAATGCTTTC CTTAGTGAGC TTAACCAT 2028

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTCTGGTCTT CTTCAAACT CA 22

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 18 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

35 CTGTAGCAT GAGCAAAG 18

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 467 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

5 (II) MOLECULE TYPE: peptide

(XI) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met	Ser	Ser	Gly	Val	Tyr	Glu	Ser	Ile	Ile	Gln	Thr	Lys	Ala	Ser	5	10	15
Val	Trp	Gly	Ser	Thr	Ala	Ser	Gly	Lys	Ser	Ile	Val	Asp	Ser	Tyr	20	25	30
Trp	Ile	Tyr	Glu	Phe	Pro	Thr	Gly	Ser	Pro	Leu	Val	Gln	Thr	Gln	35	40	45
Leu	Tyr	Ser	Asp	Ser	Arg	Ser	Lys	Ser	Ser	Phe	Gly	Tyr	Thr	Ser	50	55	60
Lys	Ile	Gly	Asp	Ile	Pro	Ala	Val	Glu	Glu	Glu	Ile	Leu	Ser	Gln	65	70	75
Asn	Val	His	Ile	Pro	Val	Phe	Asp	Asp	Ile	Asp	Phe	Ser	Ile	Asn	80	85	90
Ile	Asn	Asp	Ser	Phe	Leu	Ala	Ile	Ser	Val	Cys	Ser	Asn	Thr	Val	95	100	105
Asn	Thr	Asn	Gly	Val	Lys	His	Gln	Gly	His	Leu	Lys	Val	Leu	Ser	110	115	120
Leu	Ala	Gln	Leu	His	Pro	Phe	Glu	Pro	Val	Met	Ser	Arg	Ser	Glu	125	130	135
Ile	Ala	Ser	Arg	Phe	Arg	Leu	Gln	Glu	Glu	Asp	Ile	Ile	Pro	Asp	140	145	150
Asp	Lys	Tyr	Ile	Ser	Ala	Ala	Asn	Lys	Gly	Ser	Leu	Ser	Cys	Val	155	160	165
Lys	Glu	His	Thr	Tyr	Lys	Val	Glu	Met	Ser	His	Asn	Gln	Ala	Leu	170	175	180
Gly	Lys	Val	Asn	Val	Leu	Ser	Pro	Asn	Arg	Asn	Val	His	Glu	Trp	185	190	195
Leu	Tyr	Ser	Phe	Lys	Pro	Asn	Glu	Asn	Gln	Ile	Glu	Ser	Asn	Asn	200	205	210
Arg	Thr	Val	Asn	Ser	Leu	Ala	Val	Lys	Ser	Leu	Leu	Met	Ala	Thr	215	220	225
Glu	Asn	Asn	Ile	Met	Pro	Asn	Ser	Gln	Ala	Phe	Val	Lys	Ala	Ser	230	235	240
Thr	Asp	Ser	His	Phe	Lys	Leu	Ser	Leu	Gln	Leu	Arg	Ile	Pro	Lys	245	250	255
Val	Leu	Lys	Gln	Ile	Ala	Ile	Gln	Lys	Leu	Phe	Lys	Phe	Ala	Gly	260	265	270
Asp	Glu	Thr	Gly	Lys	Ser	Phe	Tyr	Leu	Ser	Ile	Ala	Cys	Ile	Pro	275	280	285

SUB A6
CMT

000001-000000

	Asn	His	Asn	Ser	Val	Glu	Thr	Ala	Leu	Asn	Val	Thr	Val	Ile	Cys
					290					295					300
	Arg	His	Gln	Leu	Pro	Ile	Pro	Lys	Ser	Lys	Ala	Pro	Phe	Glu	Leu
					305					310					315
5	Ser	Met	Ile	Phe	Ser	Asp	Leu	Lys	Glu	Pro	Tyr	Asn	Thr	Val	His
					320					325					330
	Asp	Pro	Ser	Tyr	Pro	Gln	Arg	Ile	Val	His	Ala	Leu	Leu	Glu	Thr
					335					340					345
	His	Thr	Ser	Phe	Ala	Gln	Val	Leu	Cys	Asn	Lys	Leu	Gln	Glu	Asp
10					350					355					360
	Val	Ile	Ile	Tyr	Thr	Ile	Asn	Ser	Pro	Glu	Leu	Thr	Pro	Ala	Lys
					365					370					375
	Leu	Asp	Leu	Gly	Glu	Arg	Thr	Leu	Asn	Tyr	Ser	Glu	Asp	Ala	Ser
					380					385					390
15	Lys	Lys	Lys	Tyr	Phe	Leu	Ser	Lys	Thr	Leu	Glu	Cys	Leu	Pro	Val
					395					400					405
	Asn	Val	Gln	Thr	Met	Ser	Tyr	Leu	Asp	Ser	Ile	Gln	Ile	Pro	Ser
					410					415					420
20	Trp	Lys	Ile	Asp	Phe	Ala	Arg	Gly	Glu	Ile	Arg	Ile	Ser	Pro	Gln
					425					430					435
	Ser	Thr	Pro	Ile	Ala	Arg	Ser	Leu	Leu	Lys	Leu	Asp	Leu	Ser	Lys
					440					445					450
	Ile	Lys	Glu	Lys	Lys	Ser	Leu	Thr	Trp	Glu	Thr	Ser	Ser	Tyr	Asp
					455					460					465
25	Leu	Glu;													

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

	(A) LENGTH:	258 amino acids
	(B) TYPE:	amino acid
30	(C) STRANDEDNESS:	single
	(D) TOPOLOGY:	linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

35	Met	Ser	Lys	Val	Lys	Leu	Thr	Lys	Glu	Asn	Ile	Val	Ser	Leu	Leu
					5					10					15
	Thr	Gln	Ser	Ala	Asp	Val	Glu	Phe	Glu	Glu	Asp	Gln	Asn	Gln	Val
					20					25					30
	Ala	Phe	Asn	Phe	Lys	Thr	Phe	Cys	Gln	Glu	Asn	Leu	Asp	Leu	Ile
					35					40					45
40	Lys	Lys	Met	Ser	Ile	Thr	Ser	Cys	Leu	Thr	Phe	Leu	Lys	Asn	Arg
					50					55					60
	Gln	Gly	Ile	Met	Lys	Val	Val	Asn	Gln	Ser	Asp	Phe	Thr	Phe	Gly
					65					70					75

SUB A6
CONT

SECRET - 68252160

	Lys	Val	Thr	Ile	Lys	Lys	Asn	Ser	Glu	Arg	Val	Gly	Ala	Lys	Asp
					80					85					90
	Met	Thr	Phe	Arg	Arg	Leu	Asp	Ser	Met	Ile	Arg	Val	Lys	Leu	Ile
					95					100					105
5	Glu	Glu	Thr	Ala	Asn	Asn	Glu	Asn	Leu	Ala	Ile	Ile	Lys	Ala	Lys
					110					115					120
	Ile	Ala	Ser	His	Pro	Leu	Val	Gln	Ala	Tyr	Gly	Leu	Pro	Leu	Ala
					125					130					135
	Asp	Ala	Lys	Ser	Val	Arg	Leu	Ala	Ile	Met	Leu	Gly	Gly	Ser	Ile
10					140					145					150
	Pro	Leu	Ile	Ala	Ser	Val	Asp	Ser	Phe	Glu	Met	Ile	Ser	Val	Val
					155					160					165
	Leu	Ala	Ile	Tyr	Gln	Asp	Ala	Lys	Tyr	Lys	Glu	Leu	Gly	Ile	Glu
					170					175					180
15	Pro	Thr	Lys	Tyr	Asn	Thr	Lys	Glu	Ala	Leu	Gly	Lys	Val	Cys	Thr
					185					190					195
	Val	Leu	Lys	Ser	Lys	Gly	Phe	Thr	Met	Asp	Asp	Ala	Gln	Ile	Asn
					200					205					210
	Lys	Gly	Lys	Glu	Tyr	Ala	Lys	Ile	Leu	Ser	Ser	Cys	Asn	Pro	Asn
20					215					220					225
	Ala	Lys	Gly	Ser	Ile	Ala	Met	Asp	Tyr	Tyr	Ser	Asp	Asn	Leu	Asp
					230					235					240
	Lys	Phe	Tyr	Glu	Met	Phe	Gly	Val	Lys	Lys	Glu	Ala	Lys	Ile	Ala
					245					250					255
25	Gly	Val	Ala												

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3049 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DAN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

	AGAGCAATTG	GGTCATTTT	TATTCTAAAT	CGAACCTCAA	CTAGCAAATC	50
35	TCAGAACTGT	AATAAGCACA	AGAGCACAAG	AGCCACAATG	TCATCAGGTG	100
	TTTATGAATC	GATCATTCAG	ACAAAGGCTT	CAGTTTGGGG	ATCGACAGCA	150
	TCTGGTAAAT	CCATCGIGGA	TTCTTACTGG	ATTTATGAGT	TTCCAACTGG	200
	TTCTCCACTG	GTTCAAATC	AGTTGTACTC	TGATTCGAGG	AGCAAAAGTA	250
	GCTTGGGCTA	CACCTCAAAA	ATTGGGIGATA	TTCCTGCTGT	AGAGGAGGAA	300
40	ATTTTATCTC	AGAACTTCA	TATCCAGTGG	TTTGATGATA	TTGATTTTCA	350
	CATCAATATC	AATGATTCTT	TCTTGGCAAT	TTCTGTTTGT	TCCAACACAG	400

SUB A6
cont

00453703 10300

	TTAACAACAA TGGAGTGAAG CATCAGGGTC ATCTTAAAGT TCCTTCTCTT	450
	GCCGAATTGC ATCCCTTTGA ACCTGTGATG AGCAGGTCAG AGATTGCTAG	500
	CASATTCGG CTCCAAGAAG AAGATATAAT TCCTGATGAC AAATATATAT	550
	CIGCTGCTAA CAAGGGATCT GTCTCTGTG TCAAAGAACA TACTTACAAA	600
5	GTGGAAATGA GGCACAATCA GGCCTTAGGC AAAGTGAATG TTCTTTCTCC	650
	TAACAGAAAT GTTCATGAGT GGCTGTATAG TTTCAAAACA AATTTCACCC	700
	AGATCGAAAG TAATAACAGA ACTGTAAATT CTCTGTCAGT CAAATCTTTG	750
	CTGATGGCTA CAGAAAAGAA CATTATGCTT AACTCTCAAG CTTTTGTATA	800
	AGCTTGTACT GATTCATATT TTAAGTTGAG CCTTGGCTG AGAATTCCAA	850
10	AAGTTTTGAA GCAAATAGCC ATACAGAAGC TCTTCAAGTT TGCAGGAGAC	900
	GAAACCGGTA AAAGTTTCTA TTTGTCTATT GCATGCATCC CAAATCACAA	950
	CAGTGTGGAA ACAGCTTTAA ATGTCACTGT TATATGTAGA CATCAGCTTC	1000
	CAATCCCTAA GTCCAAAGCT CCTTTTGAAT TATCAATGAT TTTCTCCGAT	1050
	CTGAAAGAGC CTTACAACAC TGTGCATGAT CCTTCATATC CTCAAAGGAT	1100
15	TGTTCATGCT TTGCTTGAGA CTCACACTTC CTTTGCACAA GTTCTCTGCA	1150
	ACAAGCTGCA AGAAGATGTG ATCATATATA CTATAAACAG CCTTGAACCTA	1200
	AOCOCAGCTA AGCTGGATCT AGGTGAAAGA AOCCTGAACT ACAGTGAAGA	1250
	TGCTTCGAAG AAGAAGTATT TTCTTTCAAA AACACTCGAA TGCTTGCCAG	1300
	TAAATGTGCA GACTATGTCT TATTTGCATA GCATCCAGAT TCCTTCATGG	1350
20	AAGATAGACT TTGCCAGAGG AGAGATCAGA ATCTOCCCTC AATCTACTCC	1400
	TATTTGAAGA TCTTTGCTCA AGCTGGATTY GAGCAAGATC AAGGAAAAGA	1450
	AGTCCCTGAC TTGGGAACA TOCAGCTATG ATCTAGAATA AAAGTGGCTC	1500
	ATACTACTCT AAGTAGTATT TGTCAACTTG CTTATCCTTT ATGTGTGTTA	1550
	TTTCTTTTAA ATCTAAAGTA AGTTAGATTG AAGTAGTTTA GTATGCTATA	1600
25	GCATTATTAC AAAAAATACA AAAAAATACA AAAAAATACA AAAAAATATA	1650
	AAAACCCAAA AAGATCCCAA AAGGGACGAT TTGGTTGATT TACTCTGTTT	1700
	TAGGCTTATC TAAGCTGCTT TTGTTTGAGC AAAATAACAT TGTAACATGC	1750
	AATAACTGGA ATTTAAAGTC CTAAAGAAG TTTCAAAGGA CAGCTTAGCC	1800
	AAAATTGGTT TTTGTTTTTG TTTTTFGT TTTTGTFTT TGTFTTATT	1850
30	TTTATTTTAA GTTATTTT TTGTTTTGTT ATTTTATTT TTTTFTATT	1900
	TTCFTTTATT TTATTTATAT ATATATCAAA CACAATCCAC ACAATAATT	1950
	TTAATTTCAA ACATCTACT GATTTAACAC ACTTAGCCTG ACTTTATCAC	2000
	ACTTAACACG CTTAGTTAGG CTTTAACACA CTGAACGAA TTAAAACACA	2050
	CTTAGTATTA TGCACTCTT AATTAACACA CTTTAATAAT ATGCATCTCT	2100
35	GAATCAGCCT TAAAGAAGCT TTTATGCAAC ACCAGCAATC TTGGCTCTT	2150

SUB A6
CONT

000001-000000

TCTTAACTCC AAACATTICA TAGAATTGT CAAGATTATC ACTGTAATAG 2200
 TCCATAGCAA TGCTTCGGTT AGCATTGGGA TTGCAAGAAC TAAGTATCTT 2250
 GGCATATTCT TTCCCTTTGT TTATCGIGGC ATCATCCATT GTAAATCCCTT 2300
 TGCTTTTAAG CACTGTGCAA ACCTTCCCA GAGCTTCCTT AGTGTGTIAC 2350
 5 TTAGTGGTT CAATCCGTAA CTCCTGTIAC TTTCATCTT GATATATGGC 2400
 AAGAACAACA CTGATCATCT CGAAGCTGTC AACAGAAGCA ATGAGAGGGA 2450
 TACTACCTCC AAGCATATA GCAAGICTCA CAGATTTTGC ATCTGCCAGA 2500
 GGCAGCCCGT AAGCTTGGAC CAAAGGGTGG GAGGCAATTT TTGCTTTGAT 2550
 AATAGCAAGA TTCTCATTTT TTTCAGTCTC TTCTATGAGC TTCACTCTTA 2600
 10 TCAIGCTATC AAGCCCTCTG AAAGTCATAT CCTTAGCTCC AACTCTTTCA 2650
 GAATTTTTCT TTATGTGAC CTTACCAAAA GTAAAATCAC TTGTTTCAC 2700
 AACTTTTATA ATGCCCTGGC GATTCITCAA GAAAGTCAAA CATGAAGTGA 2750
 TACTGATTTT CTTAATCAGG TCAAGATTTT CCTGACAGAA AGTCTTAAAG 2800
 TTGAATGCGA CCGGTTCCTG GTCCTCTTCA AACTCAACAT CTGCAGATTG 2850
 15 AGTTAAAAGA GAGACAATGT TTTCTTTTGT GAGCTTGACC TTAGACATGG 2900
 TGGCAGTTTA GATCTAGACC TTTCTCGAGA GATAAGATTG AAGGTGAGAA 2950
 AGTGAACAC TGTAGACCGC GGTGCTTACT TATCTGTGTA ATGTGATGAT 3000
 TTGTATTGCT GAGTATTAGG TTTTGAATA AAATTGACAC AATTGCTCT 3049

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 778 base pairs
- (B) TYPE: Nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATG CAA CAC CAG CAA TCT TGG CCT CTT TCT TAA CTC CAA 39
 ACA TTT CAT AGA ATT TGT CAA GAT TAT CAC TGT AAT AGT 78
 CCA TAG CAA TGC TTC CCT TAG CAT TGG GAT TGC AAG AAC 117
 30 TAA GTA TCT TGG CAT ATT CTT TCC CTT TGT TTA TCT GTG 156
 CAT CAT CCA TTG TAA ATC CTT TGC TTT TAA GCA CTG TGC 195
 AAA CCT TCC CCA GAG CTT CCT TAG TGT TGT ACT TAG TTG 234
 GTT GAA TCC CTA ACT CCT TGT ACT TTG CAT CTT GAT ATA 273
 TGG CAA GAA CAA CAC TGA TCA TCT CGA AGC TGT CAA CAG 312
 35 AAG CAA TGA GAG GGA TAC TAC CTC CAA GCA TTA TAG CAA 351
 GTC TCA CAG ATT TTG CAT CTG CCA GAG GCA GGC CGT AAG 390

CTT GGA CCA AAG GGT GGG AGG CAA TTT TTG CTT TGA TAA 429
 TAG GAA GAT TCT GAT TGT TTG CAG TCT CTT CTA TGA GCT 468
 TCA GTC TTA TCA TGC TAT CAA GOC TCC TGA AAG TCA TAT 507
 OCT TAG CTC CAA CTC TTT CAG AAT TTT TCT TTA TCG TGA 546
 5 OCT TAG CAA AAG TAA AAT CAC TTT GGT TCA CAA CTT TCA 585
 TAA TGG CTT GGC GAT TGT TCA AGA AAG TCA AAC ATG AAG 624
 TGA TAC TCA TTT TCT TAA TCA GGT CAA GAT TTT OCT GAC 663
 AGA AAG TCT TAA AGT TGA ATG CGA OCT GGT TCT GGT CTT 702
 CTT CAA ACT CAA CAT CTG CAG ATT GAG TTA AAA GAG AGA 741
 10 CAA TGT TTT CTT TTG TGA GCT TGA OCT TAG ACA TGG 778

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTTCGAGAT TTGCTAGT 18

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTATATCTTC TTCTTGA 18

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1401 base pairs
 (B) TYPE: Nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATG TCA TCA GGT GTT TAT GAA TCG ATC ATT CAG ACA AAG 39
 GCT TCA GTT TGG GGA TCG ACA GCA TCT GGT AAG TCC ATC 78

SUB A6
CONT0012533 " 4032900
000000 " 0000000

GTG GAT TCT TAC TGG ATT TAT GAG TTT CCA ACT GGT TCT 117
 CCA GTG GTT CAA ACT CAG TTG TAC TCT GAT TCG AGG AGC 156
 AAA AGT AGC TTC GGC TAC ACT TCA AAA ATT GGT GAT ATT 195
 OCT GCT GTA GAG GAG GAA ATT TTA TCT CAG AAC GTT CAT 234
 5 ATC CCA GTG TTT GAT GAT ATT GAT TTC AGC ATC AAT ATC 273
 AAT GAT TCT TTC TTG GCA ATT TCT GTT TGT TCC AAC ACA 312
 GTT AAC ACC AAT GGA GTG AAG CAT CAG GGT CAT CTT AAA 351
 GTT CTT TCT CTT GCC CAA TTG CAT CCC TTT GAA OCT GTG 390
 ATG AGC AGG TCA GAG ATT GCT AGC AGA TTC CGG CTC CAA 429
 10 GAA GAA GAT ATA ATT CCT GAT GAC AAA TAT ATA TGT GCT 468
 GCT AAC AAG GGA TCT CTC TCC TGT GTC AAA GAA CAT ACT 507
 TAC AAA GTC GAA ATG AGC CAC AAT CAG GCT TTA GGC AAA 546
 GTG AAT GTT CTT TCT CCT AAC AGA AAT GTT CAT GAG TGG 585
 CTG TAT AGT TTC AAA CCA AAT TTC AAC CAG ATC GAA AGT 624
 15 AAT AAC AGA ACT GTA AAT TCT CTT GCA GTC AAA TCT TTG 663
 CTC ATG GCT ACA GAA AAC AAC ATT ATG OCT AAC TCT CAA 702
 GCT TTT GTT AAA GCT TCT ACT GAT TCT CAT TTT AAG TTG 741
 AGC CTT TGG CTG AGA ATT CCA AAA GTT TTG AAG CAA ATA 780
 GGC ATA CAG AAG CTC TTC AAG TTT GCA GGA GAC GAA ACC 819
 20 GGT AAA AGT TTC TAT TTG TCT ATT GCA TGC ATC CCA AAT 858
 CAC AAC AGT GTG GAA ACA GCT TTA AAT GTC ACT GTT ATA 897
 TGT AGA CAT CAG CTT CCA ATC CCT AAG TCC AAA GCT CCT 936
 TTT GAA TTA TCA ATG ATT TTC TCC GAT CTG AAA GAG CCT 975
 TAC AAC ACT GTG CAT GAT OCT TCA TAT OCT CAA AGG ATT 1014
 25 GTT CAT GCT TTG CTT GAG ACT CAC ACT TCC TTT GCA CAA 1053
 GTT CTC TGC AAC AAG CTG CAA GAA GAT GTG ATC ATA TAT 1092
 ACT ATA AAC AGC OCT GAA CTA ACC CCA GCT AAG CTG GAT 1131
 CTA GGT GAA AGA ACC TTG AAC TAC AGT GAA GAT GCT TCG 1170
 AAG AAG AAG TAT TTT CTT TCA AAA ACA CTC GAA TGC TTG 1209
 30 CCA GTA AAT GTG CAG ACT ATG TCT TAT TTG GAT AGC ATC 1248
 CAG ATT OCT TCA TGG AAG ATA GAC TTT GCC AGA GGA GAG 1287
 ATC AGA ATC TCC OCT CAA TCT ACT CCT ATT GCA AGA TCT 1326
 TTG CTC AAG CTG GAT TTG AGC AAG ATC AAG GAA AAG AAG 1365
 TCC TTG ACT TGG GAA ACA TCC AGC TAT GAT CTA GAA 1401

35 (2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 777 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

5
 10
 15
 20
 25
 30
 35

ATG TCT AAG GTC AAG CTC ACA AAA GAA AAC ATT GTC TCT CTT TTA 45
 ACT CAA TCT GCA GAT GTT GAG TTT GAA GAA GAC CAG AAC CAG GTC 90
 GCA TTC AAC TTT AAG ACT TTC TGT CAG GAA AAT CTT GAC CTG ATT 135
 AAG AAA ATG AGT ATC ACT TCA TGT TTG ACT TTC TTG AAG AAT CGC 180
 CAA GGC ATT ATG AAA GTT GTG AAC CAA AGT GAT TTT ACT TTT GGT 225
 AAG GTC ACG ATA AAG AAA AAT TCT GAA AGA GTT GGA GCT AAG GAT 270
 ATG ACT TTC AGG AGG CTT GAT AGC ATG ATA AGA GTG AAG CTC ATA 315
 GAA GAG ACT GCA AAC AAT GAG AAT CTT GCT ATT ATC AAA GCA AAA 360
 ATT GCC TCC CAC CCT TTG GTC CAA GCT TAC GGG CTG CCT CTG GCA 405
 GAT GCA AAA TCT GTG AGA CTT GCT ATA ATG CTT GGA GGT AGT ATC 450
 CCT CTC ATT GCT TCT GTT GAC AGC TTC GAG ATG ATC AGT GTT GTT 495
 CTT GCC ATA TAT CAA GAT GCA AAG TAC AAG GAG TTA GGG ATT GAA 540
 CCA ACT AAG TAC AAC ACT AAG GAA GCT CTG GGG AAG GTT TGC ACA 585
 GTG CTT AAA AGC AAA GGA TTT ACA ATG GAT GAT GCA CAG ATA AAC 630
 AAA GGG AAA GAA TAT GCC AAG ATA CTT AGT TCT TGC AAT OCC AAT 675
 GCT AAG GGA AGC ATT GCT ATG GAC TAT TAC AGT GAT AAT CTT GAC 720
 AAA TTC TAT GAA ATG TTT GGA GTT AAG AAA GAG GCC AAG ATT GCT 765
 GGT GTT GCA TAA 777

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TACTTATCTA GAACCATGGA CAAAGCAAAG ATTACCAAGG 40

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TACAGTGGAT CCATGGTAT TTCAAATAAT TTATAAAAGC AC 42

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AGCATGGAT CCATGGTTAA CACACTAAGC AAGCAC 36

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AGCTAATCTA GAACCATGGA TGAATCACTA AGGAAAGCAT TGTTCG 46

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCCCTATCC TTGCAAGAC CC 22

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TACAGTGGAT CCATGGTTAA GGTAATOCAT AGGCTTGAC 39

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AGCTAACCAT GGTTAAGCTC ACTAAGGAAA GCATTGTTGC 40

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 46 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AGCTAATCTA GAACCATGGA TGACTCACTA AGGAAAGCAT TGTTCG 46

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AGCATTGGAT CCATGGTTAA CACACTAAGC AAGCAC 36

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TACAGTTCTA GAACCATGGA TGATGCAAG TCTGTGAGG 39

SUB A-6
cont

662201 " 00292160

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 49 base pairs
(B) TYPE: Nucleic acid.
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

AGATTCTCTA GACCATGGTG ACTTGATGAG CAAAGTCTGT GAGGCTTGC ,49

5
SUB AS
CONT

10 Thus while we have illustrated and described the preferred
embodiments of our invention, it is to be understood that this invention
is capable of variation and modification, and we therefore do not wish
to be limited to the precise terms set forth, but desire to avail
ourselves of such changes and alterations which may be made for
15 adapting the invention to various usages and conditions. Such
variations and modifications, for example, would include the
substitution of structurally similar nucleic acid sequences in which the
difference between the sequence shown and the variation sequence is
such that little if any advantages are available with the variation
20 sequence, i.e. that the sequences produce substantially similar results
as described above. Thus, changes in sequence by the substitution,
deletion, insertion or addition of nucleotides (in the nucleotide
sequences) or amino acids (in the peptide sequences) which do not
substantially alter the function of those sequences specifically
25 described above are deemed to be within the scope of the present
invention. In addition, it is our intention that the present invention may
be modified to join the N genes of various isolates that provide
resistance or immunity to *Tospovirus* infection of plants according to
the present invention into a single cassette, and to use this cassette as
30 a transgene in order to provide broad resistance to the Tospoviruses,
especially to TSWV-BL, TSWV-B, and INSV. Accordingly, such changes
and alterations are properly intended to be within the full range of
equivalents, and therefore within the purview of the following claims.

35 Having thus described our invention and the manner and a process
of making and using it in such full, clear, concise and exact terms so as

to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same;

00426703.402209